Evaluation of remote sensing based spectral and 3D point cloud data for crop biomass estimation in southern India

Dissertation to obtain the academic degree

Doktor der Agrarwissenschaften (Dr. agr.)

Submitted to Faculty of Organic Agricultural Sciences

of the University of Kassel

Supriya Dayananda

Witzenhausen, September 2019

This work has been accepted by the Faculty of Organic Agricultural Sciences of the University of Kassel as a thesis for acquiring the academic degree of Doktor der Agrarwissenschaften (Dr. agr.).

1. Supervisor: Prof. Dr. Michael Wachendorf (University of Kassel)
2. Co-supervisor: Prof. Dr. Andreas Bürkert (University of Kassel)

Day of Defence: 25.10.2019

Preface

This thesis is submitted to the Faculty of Organic Agricultural Sciences of the University of Kassel to fulfil the requirements for the degree Doktor der Agrarwissenschaften (Dr. agr.). Two papers as first and second author are published in international refereed journal (list of the original papers is given on the page V) and a third manuscript in preparation are the basis for this dissertation.

My sincere gratitude goes to my first supervisor Prof. Dr. Michael Wachendorf for his patient guidance, continuous support and valuable advice throughout the study. I highly appreciate him for being available all the time for constructive scientific discussions and for trusting me by providing the opportunity to learn and explore. I would also like to thank my second supervisor Prof. Dr. Andreas Bürkert for guiding, supporting and persuading me to continue.

I am thankful to Dr. Thomas Astor for introducing and educating me to remote sensing and for being supportive whenever I had problems under field conditions and in data analysis. I am also thankful to Dr. Thomas Fricke and his family for being so friendly and making me feel at home.

I am grateful to my scientific colleagues Jayan, Isaac, Ben, Matthias, Damian and Esther for always being there to discuss and solve problems. I always relied on them during difficult times.

My research wouldn't have been possible without the technical support of Andrea Gerke, who arranged the materials for field work and taught me the lab methodologies. I am also thankful to the field assistants Rajanna and Dhananjaya at GKVK, Bengaluru, India for their constant support in field activities and data collection.

I am thankful to each and everyone in GNR section for creating the friendly atmosphere and all the nice chats during the coffee break. I would like to thank the German Research Foundation (DFG) and Indian Department of Biotechnology (DBT) for funding the FOR2432/1 project. Also, I am grateful to the staff of University of Agricultural Sciences, GKVK, Bengaluru for supporting the field data collection with all their resources.

Last but not the least, I would like to thank my parents, family members and friends for always encouraging, supporting and believing in me during this journey.

Dedicated to my elder sister Nakshatra,

who will soon deliver her baby to this beautiful world

List of publications

Chapter 2 Moeckel, T., Dayananda, S., Nidamanuri, R. R., Nautiyal, S., Hanumaiah, N., Buerkert, A., Wachendorf, M. (2018). Estimation of vegetable crop parameter by multi-temporal UAV-borne images. *Remote Sensing*, 10(5), 805. https://doi.org/10.3390/rs10050805

Chapter 3 Dayananda, S., Astor, T., Wijesingha, J., Subbarayappa, C. T., Hanumanthappa, D. C., Mudalagiriyappa., Nidamanuri, R. R., Nautiyal, S., Wachendorf, M. (2019). Multi-temporal monsoon crop biomass estimation using hyperspectral imaging. *Remote Sensing*, 11(15), 1771. https://doi.org/10.3390/rs11151771

Tables

Figures

Abbreviations

RS	Remote sensing
PA	Precision agriculture
UAV	Unmanned aerial vehicles
RGB	Red, green, blue
SfM	Structure from motion
HS	Hyperspectral
N	Nitrogen
UAS	University of agricultural sciences
ML	Machine learning
FMB	Fresh matter biomass
LiDAR	Light detection and ranging
P	Phosphorus
K	Potassium
GCP	Ground control points
GRVI	Green-red vegetation index
DEM	Digital elevation model
RFR	Random forest regression
SVR	Support vector regression
PLSR	Partial least square regression
RMSE	Root mean square error
SD	Sampling date
NF	N fertilizer
R	Rainfed
I	Irrigated
rRMSEP	Relative root mean square error prediction

NDVI	Normalised difference vegetation index
NIR	Near-infrared
BBCH	Biologische bundesanstalt, bundessortenamt und chemische industrie
RMSEP	Root mean square error prediction
REMRI	Red edge modified ratio index
FMY	Fresh matter yield
SVM	Support vector machines
GBT	Gradient boosting trees
LAI	Leaf area index
RMSD	Root mean square difference

Table of contents

Chapter 1

Introduction

Background

Rural-urban transformation are taking place around the world. The transformation process is projected to increase from 2018-2050 mainly in Asia and Africa. In which, India, China and Nigeria will account to its major (35%) share (United Nations 2018). Bengaluru in southern India is a representative of many fast-growing cities in Asia. It is the second fastest growing megacity (India's 10 fastest growing cities 2008) in India, due to its information technology sector and moderate climatic conditions. The majority of the India's population (60%) depends on the agriculture sector for their livelihood (Arjun 2013) and, thus, the loss of agricultural land due to increasing population and expansion of urban areas is a growing problem. With arable land becoming less available, it is very important to feed the growing population by crop intensification. Appropriate application of nutrients and water play a key role in maximising crop yields.

Agriculture in India mainly depends on monsoon rainfall, surface water and ground water irrigation. The variability of monsoon rainfall is highly fluctuating and in order to grow multiple crops throughout the year, farmers are highly dependent on local irrigation systems (Ferrant et al. 2017). Irrigated crop production accompanied by fertilizer application has enabled the country with sufficient food production and is a major contributor to the green revolution (Thenkabail et al. 2009). Hence, early estimation of yield may allow better planning and forecasting of market prices and may support a higher food security based on regional, national and global demand and supply. In this context, remote sensing (RS) can be an effective tool in monitoring crop production (Aasen et al. 2015; Rouse et al. 1974) and estimating yield (Warren and Metternicht 2005). RS allows to collect data about crop production using non-destructive methods (Burkart et al. 2015) on a large scale for many fields at the same time.

As India has multiple cropping seasons in one year, agricultural systems require continuous multi-temporal remote sensing data to monitor the crop state and for crop identification (Hannerz and Lotsch 2008; Mondal et al. 2014). Evaluation of the crop state helps to identify the required amount of nutrients and water supply. Crop identification may allow to assess the effect of urbanisation

such as the changing local demand influences the cropping pattern. Precision agriculture (PA) has been widely used for two decades but its use not common among finger millet, lablab and horticultural crops yet. The small land holdings, planting density and complex plant architecture of the crops complicate PA. Sensors from terrestrial (Nidamanuri and Zbell 2011) (Dayananda et al. 2019) to airborne (Moeckel et al. 2018), and spaceborne (Panigrahy and Sharma 1997) are being used to monitor agro-environmental systems. Recently, unmanned aerial vehicles (UAV) are being used to a large extent, than satellite and airborne platforms since they are more independent from conditions such as cloud cover, revisiting time and image resolution (C. Zhang and Kovacs 2012). The flexibility of UAV to carry various sensors i.e. RGB (red, green, blue) cameras (Moeckel et al. 2018), multispectral sensors (Maimaitijiang et al. 2017) or hyperspectral (HS) sensors (Aasen and Bolten 2018) with deploy mission and spatial resolution make them more advantageous. UAV-borne RGB cameras and HS sensors are increasingly used in monitoring crop production and estimation of biomass.

UAV-borne RGB camera

UAV are considered as a promising platform for capturing detailed imagery from agricultural fields (Malambo et al. 2018). UAV RS has been successfully used in extracting detailed information on crop health (Selsam et al. 2017), crop biomass development (N. Yu et al. 2016) and crop water status (Park et al. 2017). UAV with RGB based cameras are considered as a valuable tool in the estimation of crop phenotypic characters. Plant height is an important phenological indicator of crop growth and biomass (Lati et al. 2013; Madec et al. 2017). Traditional crop height measurements are time-consuming and difficult to implement on a large scale. Therefore, crop height derived from multi-view RGB images has been recently used in barley (Bendig et al. 2014), maize (Li et al. 2017), vineyards (Weiss and Baret 2017), wheat (Schirrmann et al. 2016), sorghum (Malambo et al. 2018) and alfalfa (Parkes et al. 2016). A computer vision technique called structure from motion (SfM) is used to obtain 3D point cloud from the UAV-borne RGB images and the crop height data is derived from the point cloud (Grüner, Astor, and Wachendorf 2019; Wijesingha et al. 2019). Despite application of 3D point clouds based on UAV imagery using SfM, they are mostly used in single date or short crop growth cycles. Hence a detailed study on the entire cropping season is of utmost importance. The assessment of UAV-borne RGB imagery utilised for

extracting crop height data of vegetable crops with the aim of estimating biomass for complete cropping season is presented in chapter 2.

Hyperspectral Sensors

HS remote sensing is widely used in agricultural applications where data is collected from ground, aerial and space platforms. Mostly, HS sensors provide near-continuous narrow band spectral data from 400 nm to 2500 nm. It has been proven that spectral reflectance from the crop that is measured by HS sensor can be employed for detection of its N concentration (Vigneau et al. 2011; K. Yu et al. 2013), biomass (Yue, Feng, Jin, et al. 2018) and water stress (Krishna et al. 2019; F. Zhang and Zhou 2019). Development of HS sensors and their application in estimation of biomass has gained an increasing attention in recent years. Multi-temporal HS data was utilised to estimate crop biomass of crops such as rice (Aasen et al. 2014), wheat (Honkavaara et al. 2013; Yue et al. 2017) and maize (Wang et al. 2017). Lablab, maize and finger millet are the major crops grown in the state of Karnataka. However, no multi-temporal HS studies on estimation of biomass with different levels of nitrogen and water supply have been carried out. Multi-year data of complete crop growth cycles due to changes in agronomic practices and climatic conditions are always better for the prediction of the biomass. Hence, a multiple year data set has been used for the estimation of biomass of monsoon crops using terrestrial HS imaging during the complete growing season in the study presented in chapter 3.

Data fusion of UAV RGB imagery and hyperspectral sensors

The availability of different sensor systems enables sensor fusion i.e. the combination of different datasets from two or more sensors with different characteristics (Pohl and van Genderen 1998). Sensor data fusion may allow a more holistic interpretation of the relationship between RS data and crop parameters. Fusing data acquired with various different sensors have improved the prediction performance of vegetation parameters such as leaf area index and biomass of various crops, such as maize (Gao et al. 2013), winter wheat (Yue, Feng, Yang, et al. 2018), soybean (Maimaitijiang et al. 2017) and barley (Rischbeck et al. 2016).

Terrestrial HS and UAV-borne RGB data were used exclusively in the estimation of crop height and biomass for cereal, legume and vegetable crops in chapter 2 and 3. Hence the data fusion of

terrestrial HS and UAV-borne RGB data was used for the estimation of height and biomass of vegetable crops during the entire cropping season (see chapter 4).

Research objectives

Focussing on different aspects of multiple cropping seasons in India, experimental fields were studied in monsoon (June to October) and dry season (January to May) at University of Agricultural Sciences (UAS), Bengaluru, India. Two experimental fields rainfed and drip irrigated were maintained with varying N levels and water supply (see Chapter 9, Gallery). Lablab, maize and finger millet were grown in both experiments during the monsoon season. The vegetable crops, tomato, cabbage and eggplant were grown in a drip irrigated experiment during the dry season.

The RS measurements with UAV-borne RGB and terrestrial HS camera were collected along with the ground truth biomass samples and height measurements throughout the cropping season. The following hypothesis were addressed:

(i) Remote sensing can be used to estimate the biomass of vegetable crops using UAV.

(ii) Terrestrial HS imaging can be used to estimate the biomass of monsoon crops.

(iii) Fusion of height and spectral data may improve biomass estimation capability of vegetable crops when compared to individual sensor data (UAV-borne RGB or HS).

(iv) Biomass estimation accuracies may vary for N levels, sampling dates and water supply of the crop.

Thus, for the assessment of multi-temporal biomass in cereals, legume and vegetable crops using RS measurements, the research objectives of the thesis were:

(i) To predict crop height and biomass of vegetable crops using UAV-borne RGB imagery.

(ii) To assess monsoon crop biomass using terrestrial HS imaging.

(iii) To estimate the vegetable crop biomass using data fusion from two sensors that is UAV-borne RGB imagery and terrestrial HS imaging.

(iv) To monitor the performance of biomass estimation models at different stages of crop development and at different N levels and water supply for monsoon crops.

(v) To evaluate the variations of prediction accuracy with the machine learning (ML) methods applied to vegetable crops.

(vi) To identify important spectral bands affecting the estimation of biomass.

Chapter 2

Estimation of Vegetable Crop Parameter by Multi-temporal UAV-Borne Images

Abstract

3D point cloud analysis of imagery collected by unmanned aerial vehicles (UAV) has been shown to be a valuable tool for estimation of crop phenotypic traits, such as plant height, in several species. Spatial information about these phenotypic traits can be used to derive information about other important crop characteristics, such as fresh biomass yield, which cannot be derived directly from the point clouds. Previous approaches have often only considered single date measurements using a single point cloud derived metric for the respective trait. Furthermore, most of the studies focused on plant species with a homogenous canopy surface. The aim of this study was to assess the applicability of UAV imagery for capturing crop height information of three vegetable crops: eggplant, tomato, and cabbage with a complex vegetation canopy surface during a complete crop growth cycle to infer crop biomass. Additionally, the effect of crop development stage on the relationship between estimated crop height and field measured crop height was examined. Our study was conducted in a multi-factorial experiment at the University of Agricultural Sciences, GKVK campus, Bengaluru, India. For all crops, crop height and the biomass were measured at five dates during one crop growth cycle between February and May 2017 (average crop height was 42.5, 35.5, and 16.0 cm for eggplant, tomato, and cabbage). Using a structure from motion approach, a 3D point cloud was created for each crop and sampling date. In total, 14 crop height metrics were extracted from the point clouds. Machine learning methods were used to create prediction models for vegetable crop height. The study demonstrates that the monitoring of crop height using an UAV during an entire growing period can result in detailed and precise estimates of crop height and biomass for all three crops (R^2 ranging from 0.87 to 0.97, bias ranging from −0.66 to 0.45 cm). The effect of crop development stage on the predicted crop height was found to be substantial (e.g., median deviation increased from 1% to 20% for eggplant) influencing the strength and consistency of the relationship between point cloud metrics and crop height estimates and, thus, should be further investigated. Altogether the results of the study demonstrate that point cloud generated from UAV-based RGB imagery can be used to effectively measure vegetable crop biomass in

larger areas (relative error = 17.6%, 19.7%, and 15.2% for eggplant, tomato, and cabbage, respectively) with a similar accuracy as biomass prediction models based on measured crop height (relative error = 21.6, 18.8, and 15.2 for eggplant, tomato, and cabbage).

Introduction

The ability to non-destructively collect information about crops makes remote sensing a less time consuming and less labour-intensive tool in agricultural science than traditional methods of crop growth monitoring. Unmanned aerial vehicles (UAV) have emerged as a promising remote sensing platform to capture detailed imagery from agriculture crop fields [1]. The ability of UAVs to collect data over wide areas in flexible intervals makes them a superior tool compared to satellite or airborne remote sensing [2]. UAV systems equipped with RGB (red, green, blue) cameras are widely distributed, but systems with other cameras installed (e.g., multi-spectral cameras) are also getting increasingly available. Detailed information about crop health [3], crop biomass development [4], and crop water status [5] have been already successfully extracted from UAV remote sensing for various agriculture crops.

Collecting information about phenotypic traits, such as plant height or biomass, at high temporal resolution, is essential for many site-specific management practices or plant physiological studies [6]. Plant height is considered as a good indicator for plant growth and biomass [7,8]. However, the use of classical ground-based height measurements is time consuming and can rarely be applied non-destructively across larger areas in a repetitive manner [8]. Alternative methods based on LiDAR (light detection and ranging) [9,10], ultrasonic sensors [11], or high resolution RGB imagery [12] have been developed recently. While LiDAR sensors provide highly accurate and dense 3D point measurements of crop surfaces, they are still very expensive, and require specific expertise for handling of sensors and the subsequent analysis of the data [13]. Although less cost intensive alternatives exist (Ehlert 2010), they cannot be used to cover large areas due to their limited mobility. Ultrasonic sensors are considered as low-cost and user-friendly approach, but are often limited by their spatial resolution and their susceptibility to wind [14]. RGB image-based detection of crop height is the most recently evolving approach for many different crops, including barley [12], maize [15], vineyards [16], wheat [17], sorghum [1], or alfalfa [18]. Especially, 3D point clouds generated from UAV-borne RGB images using SfM (structure from motion) techniques offer new options for deriving crop height information [1]. Despite the successful application of 3D point

clouds based on UAV imagery and the SfM procedure, many studies have been limited to single acquisition dates or only short periods in the crop growth cycle. In addition, most of the studies concentrate on crops, which are characterized by homogenous crop surfaces. To our knowledge, no study has yet tested the application of UAV-borne RGB imagery for estimating vegetable crop height over an entire cropping season.

The aim of this study, therefore, was to assess the applicability of UAV-borne RGB imagery for capturing crop height information of three vegetable crops eggplant, tomato, and cabbage over an entire cropping season. These three crops are major vegetable species in the larger area of Bengaluru, Karnataka, India. They represent three different crop growth forms, and canopy shape pattern ranging from a less complex ball shape for cabbage, to a complex multi-layer appearance for tomato and eggplant. The specific objectives of this study therefore were (1) evaluating the ability of UAV imagery and SfM techniques to model the height of three vegetable crops over an entire crop growth cycle, (2) monitoring the effect of the crop development on the model prediction quality of crop height values using UAV imagery, (3) quantifying the effect of crop growth form on the crop height models, and (4) evaluating the applicability of the crop height models for estimating the biomass of vegetable crops.

Methods

Study Site

The study was conducted on the experimental farm of the University of Agricultural Sciences (UAS), Bengaluru (12°58′20.79″N, 77°34′50.31″E, 920 asl), India (Figure 1A), where Kandic Paleustalfs and Dystric Nitisols are the dominating soil types. The mean annual temperature is 29.2 °C, with an average precipitation of 923 mm, of which most occurs during the monsoon season from June to October [19].

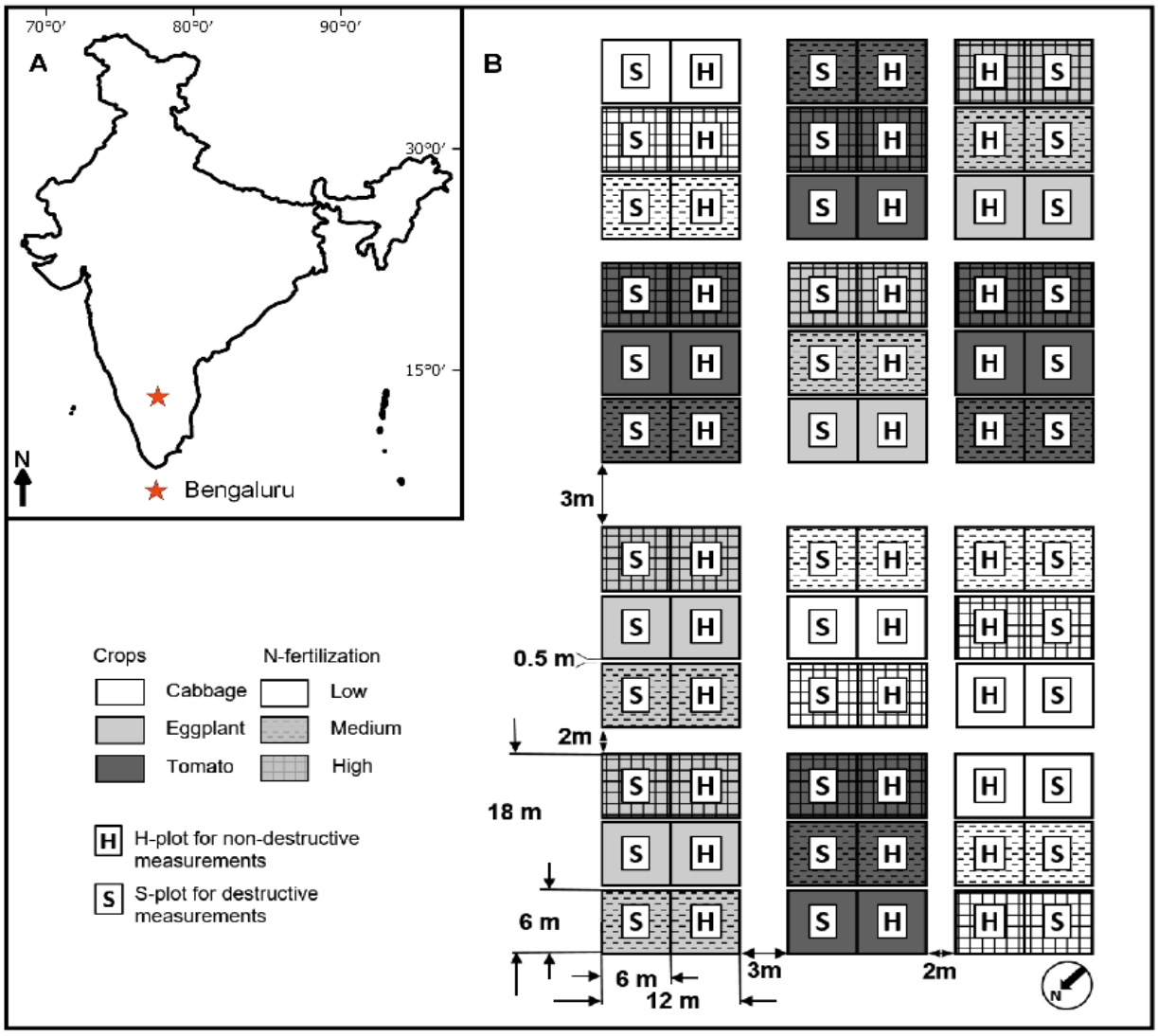


Figure 1. (A) Political map of India showing the location of Bengaluru, (B) layout of the field experiment at University of Agricultural Sciences, GKVK Campus, Bengaluru. S indicates plots in which destructively measured parameters were taken (i.e. biomass harvest). H indicates plots in which non-destructive measurements were conducted (i.e. spectral sampling). S plots were used for model calibration, while H plots were used for model validation.

Experimental Design

The field experiment was established in 2016 with a respective typical rainy and dry season crop rotation. Within the rainy season (Kharif, July to October) maize (*Zea mays* L., cultivar Nithyashree), millet (*Eleusine coracana* Gaertn., cultivar MR 65), and lablab (*Lablab purpureaus* L. Sweet, cultivar: HA 4) were grown, while in the dry season (Rabi, January to May), tomato (*Solanum lycopersicum* L., cultivar NS-501), eggplant (*Solanum melongena* L., cultivar: Ankur),

and cabbage (*Brassica oleracea* L., cultivar: Unnati) were grown on the experimental site (Figure 1B). The aim was to include the most commonly grown crops in the greater region of Bengaluru in the field experiment. The selection of crops and cultivars was based on recommendations of scientists at UAS. During the dry season, crops were irrigated with a drip irrigation system. The split-plot experiment comprised 12 main plots (12 × 18 m), whereby each treatment was replicated four times (Figure 1A). In each main plot, three subplots (6 × 12 m) with different levels of nitrogen (N) fertilizer were randomly allocated. Besides a blanket application of phosphorus and potassium (40 kg P ha^{-1} and 20, 24, and 50 kg K ha^{-1} for eggplant, tomato, and cabbage), the mean N rate (50 kg N ha^{-1} as urea for eggplant, 46 kg N ha^{-1} for tomato, and 60 kg N ha^{-1} for cabbage) reflected the recommended N dose in the region, whereas the high (+50%) and low (−50%) dose represented additional intensities typical to the study area. Nitrogen application was split into two dressings to reduce leaching losses by the typically strong rainfalls after fertilizer application. As fertilizers were not distributed evenly on the total plot area, but directly applied to individual plants, no nutrient deficiency was noticeable in either treatment. To allow for destructive measurements, each plot was divided equally in a sampling sub-subplot (S-plot; Figure 1B), which was used for all destructive biomass samplings and a harvest sub-subplot (H-plot) for undisturbed crop growth and non-destructive measurements, such as determination of plant height. The S- and H-plots had a size of 6 × 6 m each. The layout thus comprised a total of 36 subplots (3 crops × 3 fertilizer treatments × 4 replicates).

Plant Sampling and Measurements

To calculate average crop height, the height of 30 plants in each S- and H-plot was measured with a ruler to the nearest 0.1 cm (Figure 2) whereby the distance from the ground to the topmost vegetative element was determined. The measurements were conducted at five dates throughout the growing season (Table A1 and Figure A1, see Appendix). To measure total biomass (including fruits when present at later growth stages), three randomly selected plants were harvested 5 cm above ground in all S-plots (Figure 2). At each harvest, the biomass (t ha^{-1}) was weighted in the field and averaged for each plot.

RGB Imagery Sampling

Prior to destructive sampling, each main plot was scanned using a RGB camera mounted on a UAV (DJI 3 Professional, DJI, Shenzhen, Guangdong, China). The camera model used was the

standard DJI FC300X with a focal length of 3.61mm. At every sampling date, flights were conducted for each plot using the autopilot and the software Pix4D Capture (Pix4D SA, Lausanne, Switzerland). The flight path was kept constant throughout all sampling dates and for all dates. For capturing the images, a gridwise flight path was defined, for which the area of the main plot was overlaid with a grid of several perpendicular flight lines. The flight delivered nadir images, which were used to generate a 3D point cloud. On average ~200 images each with a resolution of 4000 × 3000 pixels were taken during each flight (Figure 2). The images were taken with an 80% forward and side overlap, whereby the flight altitude was set to 20 m (ranging from 17 m to 24 m, due to variations in the built-in GPS of the drone). The flying height represented a compromise of getting enough images (ranging from 189 to 219) for the point cloud processing and still covering the plot in as little time as possible. The variation in the flying height also leads to variation in the mean ground resolution ranging between 6 and 8 mm per pixel. In total, 60 flights (4 main plots × 3 crops × 5 sampling dates) were conducted. The flights were conducted only on days with no or only low wind, to reduce the effect of moving plants during recording. The spatial position of each image was measured by the internal GPS of the UAV (spatial accuracy ~2 m). However, its accuracy is not sufficient for direct georeferencing. Thus, for each main plot, four permanent ground control points (GCPs) were measured at the corners of the plot using a differential GPS (Trimble Inc., Sunnyvale, CA, USA) with a spatial accuracy of ~5 cm. These GCPs were subsequently used for georeferencing, geo-correction and co-registration of the images.

Point Cloud Processing

To receive point cloud data from the UAV images, a SfM algorithm was applied (Figure A1, Appendix). The first step of the algorithm is to extract features in each image that can be matched to their corresponding features in other images [20,21]. Prior to this matching step, the SfM performs a bundle adjustment among the images, based on matching features between the overlapped images to estimate interior and exterior orientation of the onboard sensor. Applying a multi-view stereo matching algorithm to the images, a dense point cloud is generated [22,23]. All images were processed using the software Photoscan Pro (AgiSoft LLC, St. Petersburg, Russia). The bundle adjustment was conducted with the high accuracy settings and a key point limit of 100,000 and a tie point limit of 1000. The resulting sparse point cloud was georeferenced using the GCPs to ensure precise alignment and georeferencing of the point cloud. The RMS of the projections ranged

from 0.2 to 0.4. All GCPs were manually identified in the raw images, and their coordinates were fed into the Photoscan software. The spatial error of the GCP for all datasets ranged from 5–20 cm. The accuracy decreased towards the end of growing season, due to difficulties in identifying the GCP due to overgrowing (Figure 2). Subsequently, the dense point cloud was processed using the "medium quality" setting. Although higher quality settings were possible, the medium quality was selected to reduce processing time. Depth filtering allows removing outliers within the point cloud and was set to "medium" following [24]. Afterwards, the point clouds were divided into corresponding S- and H-plot.

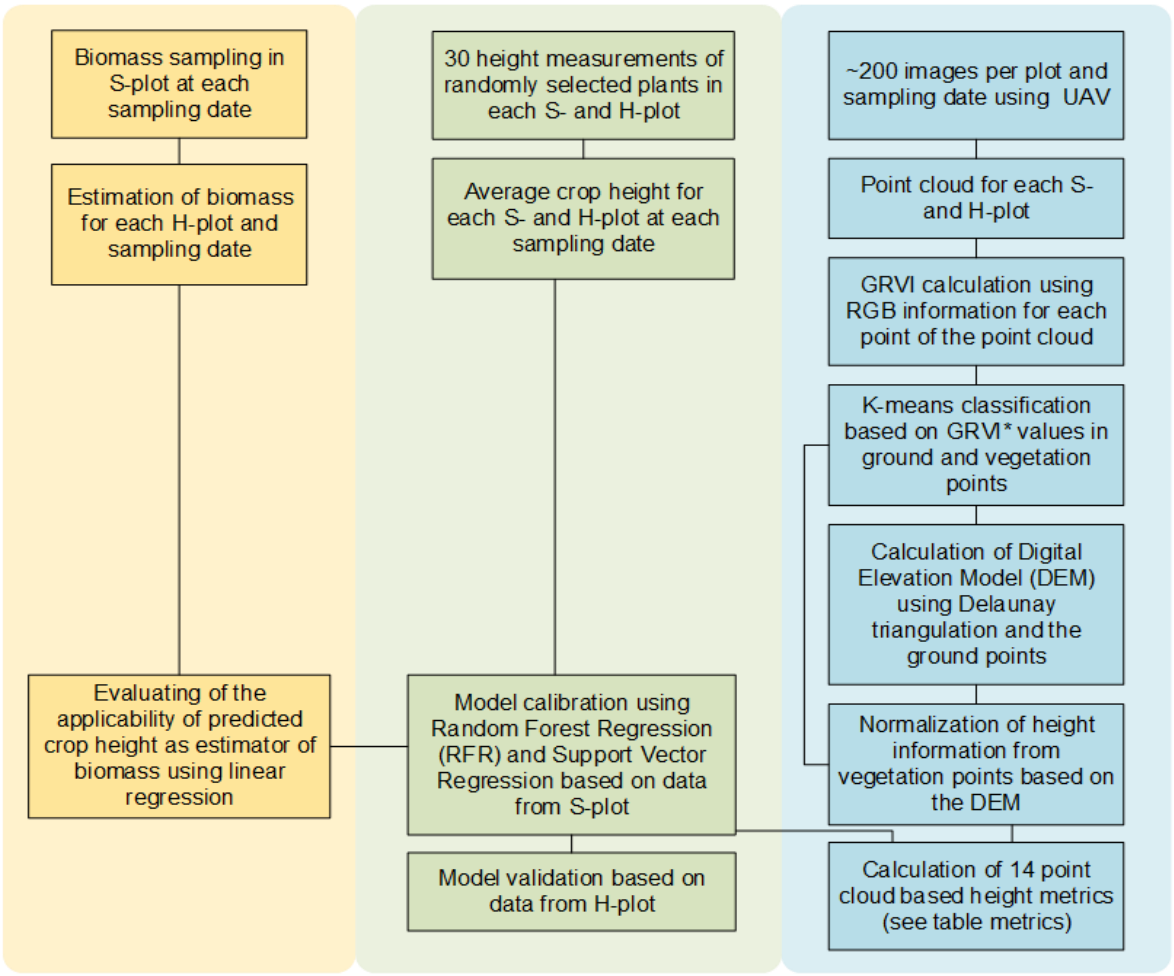


Figure 2. Process chain of the analysis, the yellow box (left) describes the biomass sampling and modelling, the green box (centre) shows the crop height sampling, and the blue box (right) describes the point cloud processing. S-plot indicates plots in which destructively measured parameters were sampled (i.e. biomass harvest). H indicates plots in which only non-destructive measurements were conducted (i.e., spectral sampling). S plots were used for model calibration, while H plots were used for model validation. * Green-Red Vegetation Index.

Ground Classification

Besides the x-, y-, z-coordinates, each point is associated with a red, green, and blue color information. This color information was used to calculate the Green-Red Vegetation Index (GRVI, Equation (1); [25]). The GRVI has been shown to be efficient in spectrally discriminating vegetation from soil [26].

$$\mathrm{GRVI} = \frac{(\mathrm{Green} - \mathrm{Red})}{(\mathrm{Green} + \mathrm{Red})}$$

with Green and Red indicating the green and red color information for all points of the point cloud. Low GRVI values are characterized by small differences between Green and Red colors, which is typical for soil. Vegetation, on the other hand, is characterized by larger differences between the Green and Red color values, which would lead to higher GRVI values. Using an unsupervised k-means algorithm, all points were classified into vegetation or ground, based on the GRVI values. For this, the k-means algorithm classified all points into k classes (here, k = 2) based on the smallest distance of the points to the mean of the class centroid. This classification was done for all plots and sampling dates individually. The classified ground points were subsequently used for the creation of a digital elevation model (DEM) using Delaunay triangulation (Figure 2). The DEM had a resolution of 1 m, which was a compromise between available ground points and necessary interpolation. To normalize the height information of the vegetation points, from each z-value, the corresponding DEM value was subtracted. This allowed for calculation of the height of each vegetation point above ground surface (Figure 2). The resulting average point density of the vegetation point cloud was 955 points m−2 and 988 points m−2 for the S- and H-plot, respectively. As plant individuals were continuously removed from the S-plot during the growing season, the total standing biomass in the S-plot steadily decreased. This, however, did not affect the presented approach, as the non-vegetated areas were removed using the GRVI values from the drone imagery, and so the average crop height was only based on actual existing plant individuals. For each plot and sampling date, 14 different point cloud metrics were calculated (Table 1). For each height metric only, the points classified as vegetation were used for calculation (Table A2, see Appendix). The selection of the metrics was based on the results from studies focusing on forest biomass estimation [27,28]. The extraction and handling of the point cloud was done using the lidR package [29] of the program R [30].

Table 1. Height metrics derived from point clouds and used as explanatory variables in modelling crop height in a multi-factorial field experiment at University of Agricultural Sciences, GKVK Campus, Bengaluru, India.

Metric	Description
Hmin	Minimum crop height
Hmax	Maximum crop height
Hmean	Mean crop height
Hsd	Standard deviation of crop height
Hmedian	Median crop height
Hskew	Skewness of crop height
Hkurt	Kurtosis of crop height
Hcv	Coefficient of variation crop height
Hq70	70th percentile of crop height
Hq80	80th percentile of crop height
Hq90	90th percentile of crop height
Hq95	95th percentile of crop height
Hq99	99th percentile of crop height
Hrelief	Crop canopy relief height (Hmean-Hmin)/(Hmax-Hmin)

Statistical Methods

In a first step, the effect of fertilizer treatment and sampling date on the measured crop height was tested using ANOVA. Prior to the ANOVA, the values were tested for normality of residuals and homoscedasticity, and if necessary, were transformed. A significant effect of the fertilizer would imply that the crop height modelling should be done separately for each N level. To predict crop height, two machine learning methods were used, namely random forest regression (RFR) [31] and support vector regression (SVR) [32]. Both methods were chosen because they can handle the high intercorrelation of the point cloud metrics (Table 1). Partial least squares regression (PLSR), another common machine learning method, was also tested for their predictive abilities. However, PLSR did not outperform RFR and SVR, and was thus removed from the further presentations. For model calibration, only point clouds from the S-plot were used, while for validation of the

model prediction quality, the point clouds from the H-plot were used. The RFR and SVR calculations were done using the packages randomForest [33] and e1071 [34] in R [30].

To evaluate model prediction quality, the explained variance using the pseudo-R2 (Equation (2) was used, as well as the root mean square error (RMSE) of the predicted values, that is, values from the H-plot (Equation (3). Additionally, the bias was calculated (Equation (4). To evaluate the prediction performance for each sampling date, the deviation of the predicted values from the measured height values, scaled by the mean of the measured height values, were used. For the comparison of the models, also, the relative error (rel.err.) was calculated (Equation (5). To assess the applicability of the predicted crop height as an estimator of biomass, a linear regression was calculated between the predicted crop height and the biomass for the H-plot. The model prediction quality was evaluated using the R2 and RMSE.

$$Pseudo - R^2 = \left[1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y}_i)^2}\right]$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{n}}$$

$$bias = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)$$

$$rel.err. = \frac{RMSE}{\max(y_i) - \min(y_i)}$$

with y_i being the measured crop height values, and $\hat{y}_i$ the predicted crop height values. $\bar{y}_i$ indicates the average measured crop height, and n is the number of samples.

Results

To test for the combined effects of sampling date and N-fertilizer on the average crop height, a two-factorial ANOVA was conducted (Table 2). For all three crops, the interaction term of sampling date and N fertilizer was not significant, while the effect of sampling date was highly significant. This is not surprising, as it was assumed that crop height increases similarly with increasing crop age at all N levels, whereby as stated earlier, the range of occurring N supply to crops was limited.

Table 2. Effects of sampling date and N fertilizer on average crop height for eggplant, tomato, and cabbage grown in a multi-factorial field experiment at University of Agricultural Sciences, GKVK Campus, Bengaluru, India.

	***p*-value**		
	Sampling Date (SD)	**N Fertilizer (NF)**	**SD × NF**
Eggplant	<0.001	0.141	0.453
Tomato	<0.001	0.978	0.720
Cabbage	<0.001	0.454	0.691

Crop Height Estimation

To predict mean crop height, 14 point cloud-derived crop height metrics were used (Table 3). The correlation among the metrics ranged from 0.01 to 0.98, with an average intercorrelation of 0.45, indicating a moderate intercorrelation (Table 3). Thus, regression methods, which can handle intercorrelations among the explanatory variables, like RFR or SVR, should be preferred to classical ordinary least squares regression models.

Table 3. Correlation among the point cloud derived crop height metrics in a multi-factorial field experiment at University of Agricultural Sciences, GKVK Campus, Bengaluru, India.

	Hmin	**Hmax**	**Hmean**	**Hsd**	**Hmedian**	**Hskew**	**Hkurt**	**Hcv**	**Hq70**	**Hq80**	**Hq90**	**Hq95**	**Hq99**
Hmax	0.01												
Hmean	0.38	0.36											
Hsd	0.1	0.67	0.75										
Hmedian	0.53	0.15	0.83	0.33									
Hskew	−0.06	0.64	−0.19	0.13	−0.25								
Hkurt	−0.02	0.55	−0.08	0.07	-0.09	0.84							
Hcv	−0.04	0.82	0.01	0.42	-0.14	0.7	0.45						
Hq70	0.24	0.43	0.97	0.86	0.67	−0.14	−0.06	0.08					
Hq80	0.18	0.43	0.91	0.91	0.53	−0.1	−0.05	0.11	0.98				
Hq90	0.15	0.49	0.86	0.95	0.45	−0.05	−0.04	0.19	0.95	0.98			
Hq95	0.12	0.55	0.79	0.97	0.39	0.01	−0.02	0.27	0.89	0.91	0.97		
Hq99	0.07	0.74	0.62	0.92	0.29	0.23	0.1	0.51	0.71	0.73	0.81	0.9	
Hrelief	0.34	−0.33	0.38	−0.04	0.57	−0.57	−0.3	−0.35	0.24	0.16	0.08	0.01	−0.14

The measured average crop height values versus the predicted crop height values are shown for eggplant, tomato, and cabbage (Figure 3). Two machine learning methods were tested: (top) ran-

dom forest regression, and (bottom) support vector regression. For RFR, the best model was acquired for cabbage (pseudo-R2 = 0.97, RMSE = 1.3, bias = 0.01). For eggplant and tomato the results were similarly good (eggplant: pseudo-R2 = 0.93, RMSE = 6.86, bias = 0.14, tomato: pseudo-R2 = 0.89, RMSE = 5.49, bias = 0.45) (Figure 3). For SVR, the results were always weaker than for RFR, with a pseudo-R2 for eggplant, tomato, and cabbage of 0.91, 0.87, and 0.89, respectively. The RMSE and the bias, on the other hand, showed larger errors for the SVR models (eggplant: RMSE = 7.36, bias = −0.66, tomato: RMSE = 5.91, bias = 0.91, and cabbage: RMSE = 2.31, bias = −0.24).

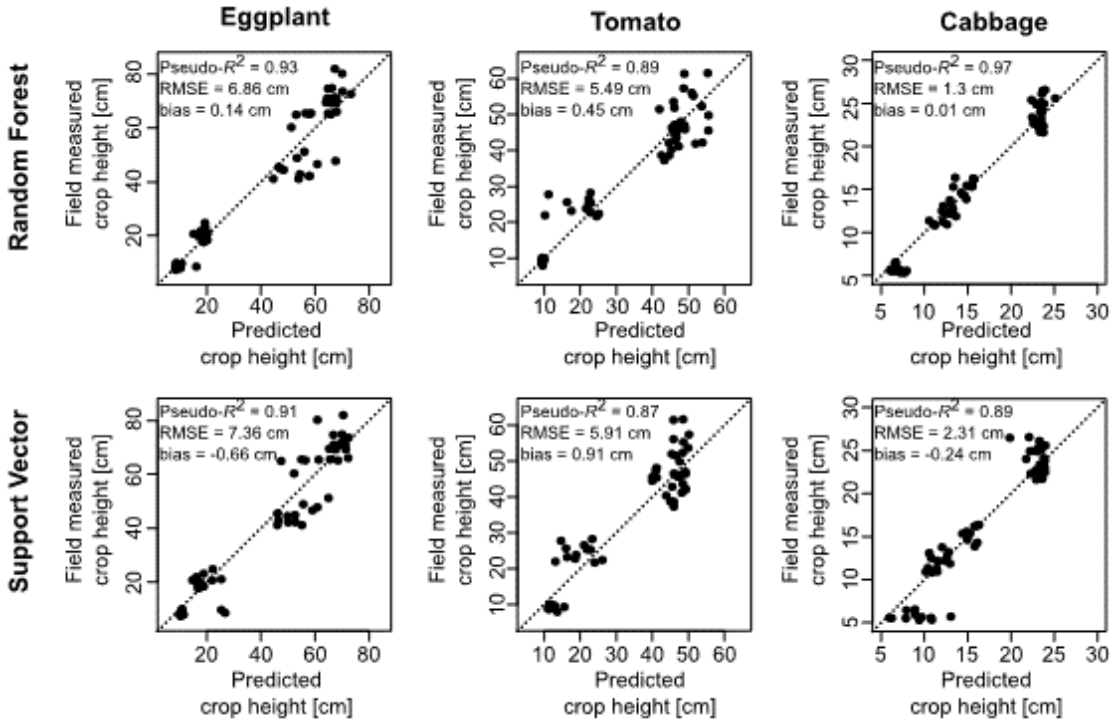


Figure 3. Field measured average crop height versus predicted average crop height for random forest regression (top) and support vector regression (bottom). From the left to the right, the results are presented for eggplant, tomato, and cabbage.

Crop Height Deviation within the Growing Season

To check whether the predictability of the average crop height varied between RFR and SVR along the growing season, the deviation of the predicted from the measured crop height was calculated (Figure 4). For eggplant the median deviation was between 1–20% for all sampling dates for both RFR and SVR. An increase of the relative deviation was visible from sampling date 1 to 5, except

for sampling date 3, for which both methods overestimated the measured average crop height. For tomato, the relative deviation shows no clear pattern along the sampling dates for both RFR and SVR. The median deviation for RFR was lowest for sampling date 1 (−0.009) and highest for sampling date 4 (0.16). For SVR, the fluctuation shows a higher amplitude than for RFR, ranging from 0.06 to −0.19. For cabbage, the deviation was lowest for RFR across all sampling dates, ranging from −0.008 in sampling date 2, to 0.10 for sampling date 5. For SVR, the highest deviation was found for sampling date 1 (−0.2). For the other sampling dates, the deviation was comparable to the results from the RFR.

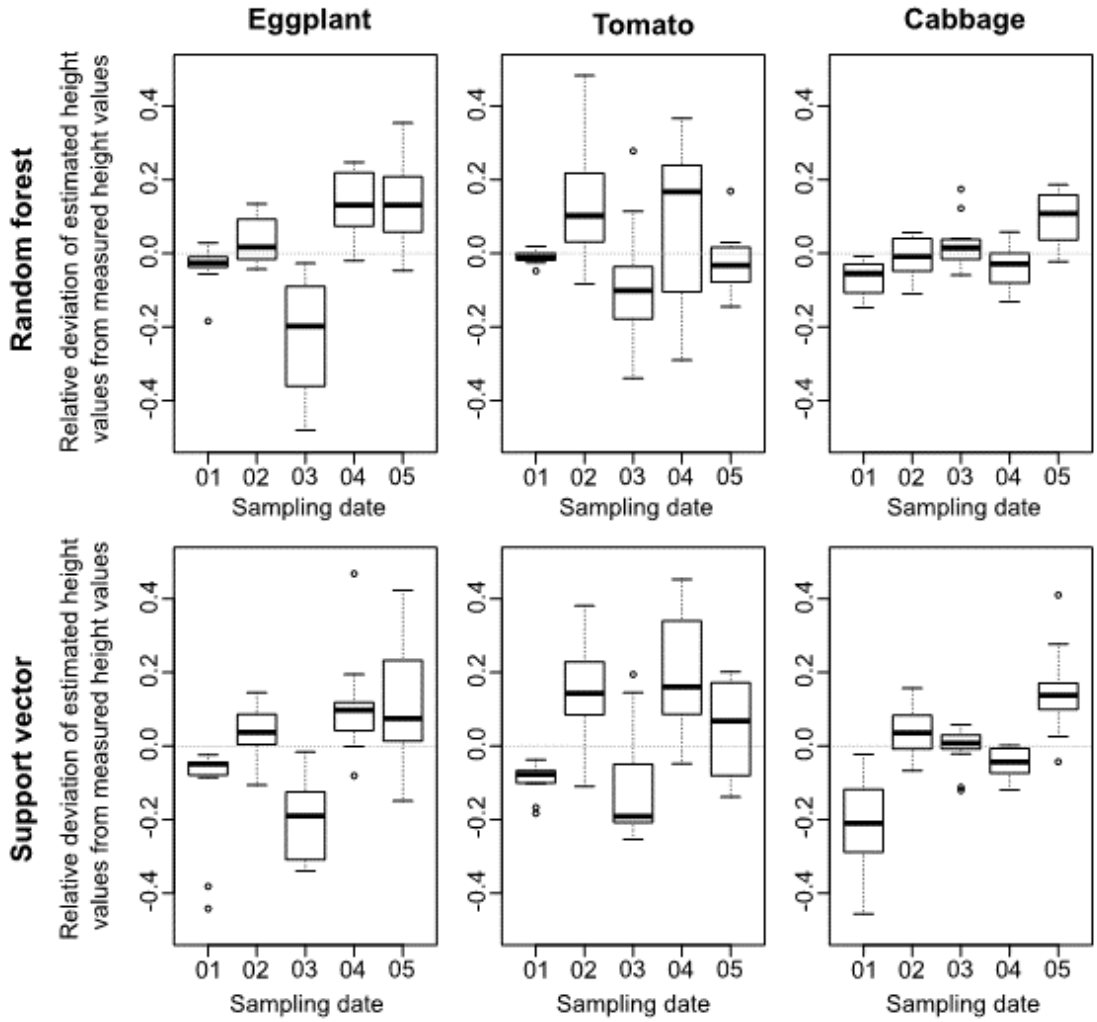


Figure 4. Relative deviation of the predicted crop height values based on random forest regression (top) and support vector regression (bottom) from the measured crop height values for each sampling date (01-05). From the left to the right, the results are presented for eggplant, tomato, and cabbage.

Crop Biomass Estimation

To test whether the predicted crop height values are reliable estimators for the biomass of the three crops, a log-linear model was calibrated with the predicted crop height values as predictors and the biomass of the H-plot as dependent variable (Figure 5). The predicted crop height values were based on RFR. The models for all three crops showed significant relationships with R^2 values from 0.90 (RMSE = 6.01 t ha−1) for eggplant, 0.88 (RMSE = 3.76 t ha−1) for tomato, and 0.95 (RMSE = 9.97 t ha−1) for cabbage. In order to compare the model based on the predicted crop height with a model based on the measured crop height, a log-linear model was also calibrated with the measured average crop height (Figures 5 and 6). For eggplant, the model on the measured crop height showed a higher relative error (rel.err. = 21.6%) than the model based on the predicted crop height. For tomato the model based on the measured crop heights showed a slightly lower error of 18.3% and for cabbage the relative error was the same (rel.err. = 15.2%) for both models.

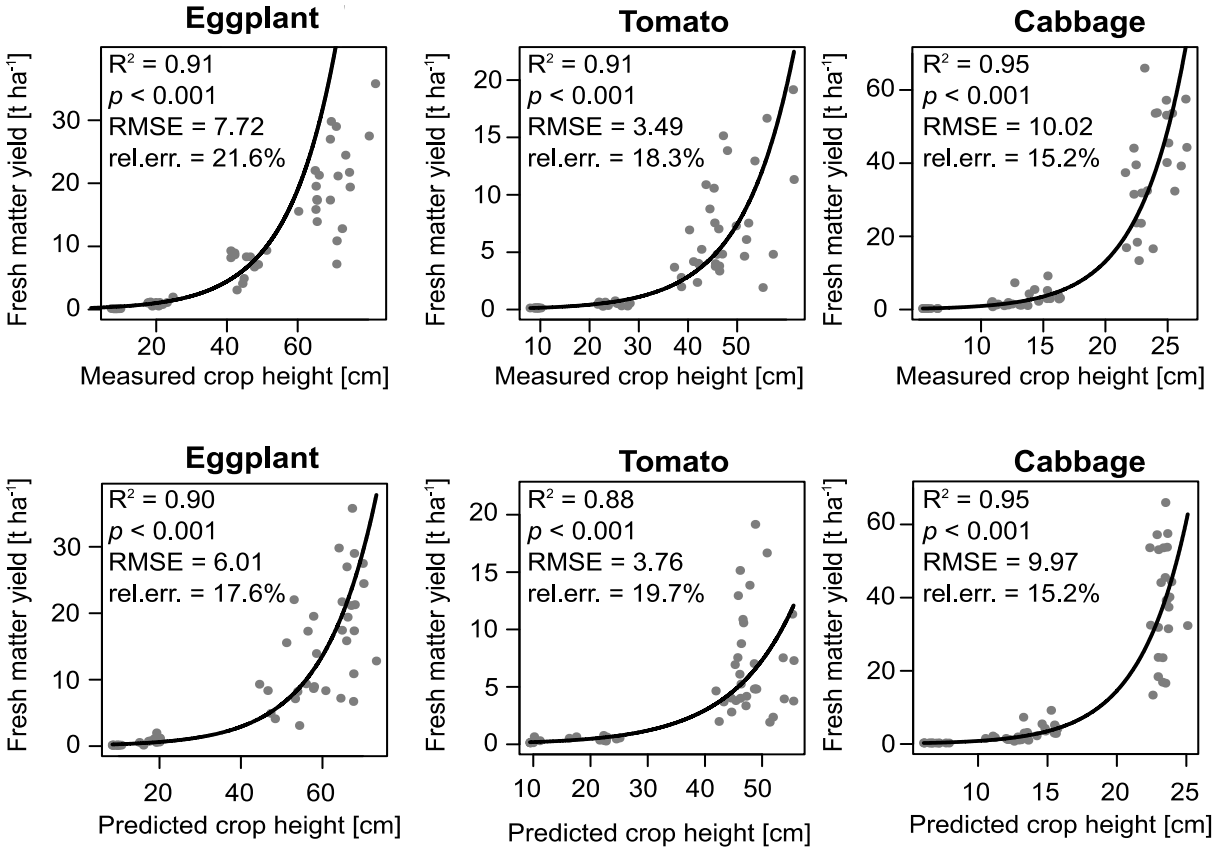


Figure 5. Biomass versus the (top) field measured crop height and (below) predicted crop height based on random forest regression.

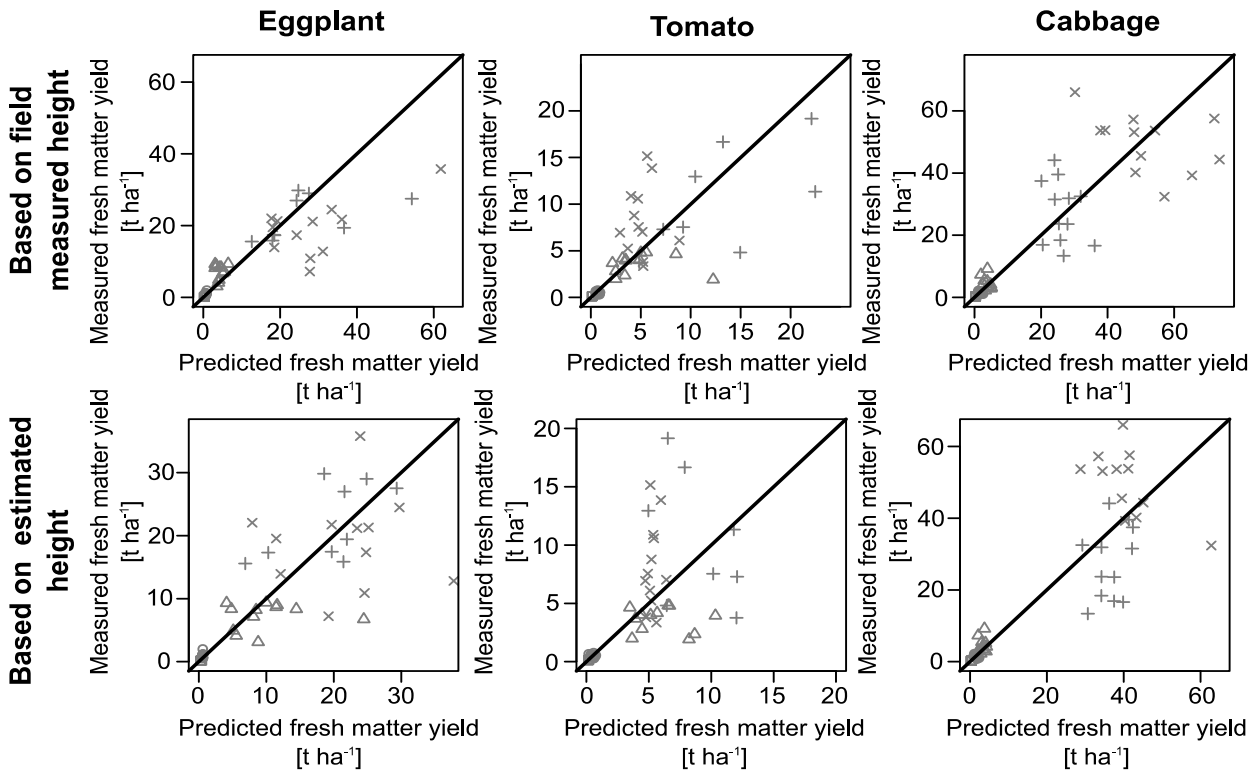


Figure 6. Predicted versus observed biomass values based on (top) field measured crop height and (bottom) predicted crop height. The symbols indicate the five sampling dates: (square) sampling date 1, (circle) sampling date 2, (triangle) sampling date 3, (plus) sampling date 4, and (cross) sampling date 5. The black lines indicate the 1:1 fit of the values.

Discussion

The aim of the current study was to assess the applicability of UAV-borne RGB imagery for capturing crop height information for the three vegetable crops (eggplant, tomato, and cabbage), and to assess whether the developed methods can be used over an entire cropping season. Overall, results indicate that a successful estimation of crop height for vegetable crops with very different crop growth forms could be achieved (with R^2 value ranging from 0.89 for tomato to 0.97 for cabbage). These results are similar to multi-temporal crop height models for maize and sorghum (ranging from 0.68 to 0.78) [1], although, in contrast to this study, no systematic bias according to the vegetable growth stage could be found.

Our experiment shows that UAV-borne RGB imagery in combination with SfM techniques can be used to estimate vegetable crop height during the complete growing period. The suggested method using several extracted crop height metrics from 3D point data was used previously in forestry

research [27], and in agroecological research [35]. For example, [35] used a group of point cloud derived height metrics to estimate variations of maize crop height at landscape scale. The advantage of this method in comparison to the more frequently used single best metric approach is that the information content about crop height differs for each single metric, as can be seen in the correlation among all metrics (Table 3). This indicates that multiple crop height metrics predict biomass information more comprehensively than single crop height metrics. Modern machine learning methods can handle highly intercorrelated variables [36] better than ordinary least squares regression models, and should therefore be the tool of choice for the prediction of crop height based on multiple crop height metrics from UAV-borne RGB imagery. The results showed no clear superiority of RFR over SVR, although RFR always yielded slightly better results (Figure 3). The slightly better performance might be explained by the lower sensitivity to data skewness and to model overfitting [37].

UAV-borne RGB imagery-based crop height values indicate a systematic deviation from the field measured crop height values (Figure 4). The apparent increase in the median bias, from early growing stages to the late stages for eggplant and cabbage, indicates an increasing overestimation of the vegetable crop height by the prediction model. A similar trend is evident in [1] for maize and sorghum. They state that the deviations are mainly due to time lags between the UAV-based measurements and the field-based measurements. For the current study, the sampling date was identical for both UAV and field measurements, indicating that other factors might be the reason for the deviation. The biased estimation is more likely a result from inaccuracies in the point cloud generation process and is, thus, likely to be related to a biomass increase during the crop development, which could complicate the SfM processing [38]. For tomato, this trend in bias could not be found, but in contrast to the other two crops tomato was tied up to horizontal threads from sampling date 3, to prevent molding of the fruits. This agricultural practice might also be the reason for the high bias values for tomato at sampling date 4 (Figure 4). In addition to the deviations in estimated crop height from the SfM algorithm, the field-based height measurements are not bias-free. For the present study, the maximum height of any vegetation part at each measurement point was used for the height measurements, which could lead to an overestimated average plot crop height.

The three vegetable crops examined in the present study represent different growth forms with different internal vegetation structures. While cabbage is characterized by a relatively flat and

round shape with compact structure, tomato and eggplants are growing taller, which is characteristic for more complex structures. In comparison to tomato, eggplant has larger leaves and thicker stems. These vegetation characteristics lead to differences in the biomass and vegetation structure, particularly at the end of the growing season (Figure 5), and could have affected the vegetable crop height estimation [39]. Despite these variations, no difference in model prediction quality was found (Figures 3 and 4). These results agree with the results of [39], who have shown that, with high spatial resolution UAV-borne RGB imagery, even small structural details can be successfully differentiated.

Crop height is considered as an important indicator for biomass of crops such as maize [35], barley [12], sorghum [1], and poppy [40]. The predicted vegetable crop height values of the present study show strong and highly significant relationships to the biomass for all three crops (Figure 5). The model prediction quality (measured as R^2) was better in the present study than in the study by [36], who estimated maize aboveground biomass using UAV-borne RGB imagery. However, in addition to being only a single date data [35], the flying height of the UAV used by [35] was much higher (130 m) than the flying height (20m) maintained in this study, leading to manifold differences in the spatial resolution of imagery. In a similar study, [41] found a comparably strong relationship between UAV-borne RGB derived from multi-temporal crop height and barley dry matter yield ($R^2 = 0.85$). The accuracies of the models based on the predicted crop height showed similar prediction qualities than the corresponding models based on field-based crop height values (Figure 5). This supports the idea that point cloud analysis could replace intensive field campaigns for measuring crop height with a similar prediction accuracy for the crop biomass. Both modelling approaches showed clear deviations for the late growing stages (growing stage 4 and 5; Figure 6). These deviations indicate that height is a less accurate estimator of biomass for crops during late growing stages. The results of the present study support the finding of other studies that high resolution multi-temporal crop height information is needed for successful estimation of crop biomass [10,41]. Alternative plant phenotypic traits, such as plant volume, may even yield more accurate predictors for biomass [40]. Thus, future studies should evaluate the further potential of 3D point cloud analysis for estimating key plant phenotypic traits, which can be used as reliable predictors for plant biomass. Furthermore, the results indicate that for a successful estimation of biomass, height information alone is not enough. Evolving sensor fusion approaches might improve the model prediction performance, as it has been shown for example for grassland biomass [42,43].

Uncertainties, Errors, and Accuracies

Disturbances of vegetation canopy through environmental conditions (e.g., wind, clouds) can affect point cloud creation and lead to inaccuracies during the point cloud processing. While wind can move the vegetation canopy and, thus, may lead to a distortion of height, clouds can lead to differences in the RGB values which were used for ground classification. However, as the average canopy height was used, we assume these effects are minor, and no UAV flight was conducted during high wind conditions.

For a successful height estimation, a highly precise georeferencing is mandatory. Four GCPs (as used in this study) represent the minimum number of control points for georeferencing. More GCPs could have led to a more accurate representation of the actual crop height. To keep the relative error between the measurements constant, permanent GCPs were used.

Accurate estimation of height for each vegetation point needs an accurate DEM model. While at the beginning of the growing season, plenty of space among individual plants allows for identifying enough ground points for an accurate calculation of the DEM, much of the ground is covered by vegetation at later crop growth stages, which reduces the amount of ground-classified points. This problem is not prominent for cabbage, as there was always enough space between the plant individuals, whereas tomato and eggplant developed a dense vegetation cover with progressing growing season.

Although the accuracies of the biomass prediction models based on the estimated crop height were similar to the predicted biomass based on the measured crop height, circa 10% of the variation in the biomass values remained unexplained. The relationship between height and biomass is exponential, indicating that at this stage large values of vegetation height are not a stable estimator of biomass anymore. A reason for this might be that the linear relationship between biomass and crop height is only valid at the beginning of the growing season, whereas with progressing growing season biomass further increases, which is rather due to the growth of fruits than due to height increase.

Conclusions

High resolution subplot level estimation of crop phenotypic traits from UAV-borne RGB imagery is fast emerging as one of the time and cost effective remote sensing tool for agricultural crops.

The processing and analysis approach developed in this study, based on the extraction of several height parameter from the 3D point clouds and machine learning regression methods, exhibit stable model prediction quality for biomass during the entire growing period and for the three vegetable crops eggplant, tomato, and cabbage. However, the results also show that crop height information, based on UAV imagery analysis, is affected by vegetable crop internal structure and agricultural management. Further research is needed to examine the effect of other crop phenotypic traits on model prediction quality. Overall, the study demonstrated that time-consuming manual height measurement can be replaced by remote sensing approaches for field vegetable crops, and recommend further multi-site and multi-crop studies to better understand the temporal variation between field measured crop height and crop height, based on UAV-borne RGB-based 3D point clouds. In the future, the suggested approach could be used for a real-time evaluation of vegetable biomass.

Appendix A

Table A1. Sampling dates for studied crops in a multi-factorial field experiment at University of Agricultural Sciences, GKVK Campus, Bengaluru, India.

	Eggplant	**Tomato**	**Cabbage**
Sampling date 1	08.03.2017	09.03.2017	07.03.2017
Sampling date 2	29.03.2017	30.03.2017	28.03.2017
Sampling date 3	20.04.2017	18.04.2017	10.04.2017
Sampling date 4	16.05.2017	04.05.2017	11.05.2017
Sampling date 5	13.06.2017	05.06.2017	07.06.2017

Appendix B

Table A2. Values of mean and standard deviation (brackets) for parameter used in the present study. The values for the height metrics and the measured canopy height are based on the values from the S- and H-plot (n = 24). The biomass information is based on the values from the H-plot (n = 12) in a multi-factorial field experiment at University of Agricultural Sciences, GKVK Campus, Bengaluru, India.

Crop	Sampling date	Measured Crop height	Biomass [kg m^{-2}]	Hmin	Hmax	Hmean	Hsd	Hmedian	Hskew	Hkurt	Hcv	Hq70	Hq80	Hq90	Hq95	Hq99	Hrelief
Cabbage	1	5.856 (0.379)	0.028 (0.028)	0.001 (0)	0.173 (0.417)	0.017 (0.050)	0.028 (0.107)	0.005 (0.001)	3.669 (3.537)	46.453 (97.946)	0.989 (0.437)	0.042 (0.145)	0.066 (0.248)	0.081 (0.306)	0.089 (0.316)	0.109 (0.329)	0.099 (0.056)
	2	11.936 (1.020)	0.139 (0.139)	0.001 (0)	0.168 (0.508)	0.012 (0.012)	0.02 (0.057)	0.007 (0.001)	1.742 (0.673)	3.857 (3.906)	1.005 (0.688)	0.026 (0.039)	0.034 (0.056)	0.051 (0.116)	0.080 (0.234)	0.108 (0.328)	0.132 (0.035)
	3	14.947 (1.061)	0.446 (0.446)	0.001 (0)	0.198 (0.446)	0.017 (0.007)	0.021 (0.042)	0.013 (0.001)	1.685 (1.356)	6.123 (12.136)	0.933 (0.737)	0.035 (0.022)	0.042 (0.034)	0.055 (0.063)	0.075 (0.130)	0.132 (0.322)	0.153 (0.052)
	4	22.296 (0.892)	2.746 (2.746)	0.001 (0)	0.235 (0.374)	0.035 (0.005)	0.032 (0.031)	0.029 (0.002)	1.132 (1.417)	2.738 (12.325)	0.854 (0.462)	0.070 (0.016)	0.080 (0.024)	0.095 (0.044)	0.114 (0.086)	0.177 (0.298)	0.207 (0.047)
	5	24.967 (1.029)	4.972 (4.972)	0.001 (0)	0.320 (0.643)	0.035 (0.004)	0.030 (0.022)	0.029 (0.002)	1.283 (1.885)	5.397 (22.835)	0.825 (0.401)	0.069 (0.011)	0.078 (0.015)	0.093 (0.027)	0.108 (0.052)	0.166 (0.233)	0.186 (0.054)
Eggplant	1	8.744 (0.883)	0.015 (0.015)	0.001 (0)	0.445 (1.299)	0.006 (0.002)	0.014 (0.033)	0.005 (0.001)	4.546 (6.417)	75.084 (171.526)	1.631 (2.860)	0.011 (0.005)	0.013 (0.007)	0.018 (0.016)	0.021 (0.017)	0.076 (0.174)	0.095 (0.056)
	2	20.020 (1.736)	0.099 (0.099)	0.001 (0)	0.100 (0.013)	0.018 (0.004)	0.016 (0.003)	0.012 (0.003)	1.321 (0.248)	1.451 (1.016)	0.928 (0.067)	0.038 (0.007)	0.044 (0.008)	0.053 (0.009)	0.060 (0.009)	0.074 (0.009)	0.168 (0.030)
	3	44.833 (3.433)	0.727 (0.727)	0.001 (0)	0.456 (0.260)	0.042 (0.004)	0.034 (0.003)	0.035 (0.004)	1.450 (0.493)	6.031 (9.016)	0.812 (0.038)	0.084 (0.008)	0.095 (0.009)	0.114 (0.010)	0.132 (0.010)	0.176 (0.016)	0.112 (0.047)
	4	68.798 (4.790)	2.239 (2.239)	0.001 (0)	0.326 (0.103)	0.074 (0.039)	0.049 (0.015)	0.066 (0.042)	0.877 (0.395)	0.894 (1.329)	0.719 (0.108)	0.134 (0.056)	0.149 (0.058)	0.171 (0.061)	0.191 (0.062)	0.229 (0.063)	0.226 (0.097)
	5	70.315 (4.778)	1.902 (1.902)	0.001 (0)	0.271 (0.042)	0.054 (0.010)	0.042 (0.007)	0.045 (0.009)	1.100 (0.167)	1.185 (0.719)	0.786 (0.036)	0.107 (0.019)	0.121 (0.021)	0.144 (0.024)	0.164 (0.026)	0.204 (0.030)	0.199 (0.033)
Crop	**Sampling date**	**Measured crop height**	**Biomass [kg m^{-2}]**	**Hmin**	**Hmax**	**Hmean**	**Hsd**	**Hmedian**	**Hskew**	**Hkurt**	**Hcv**	**Hq70**	**Hq80**	**Hq90**	**Hq95**	**Hq99**	**Hrelief**
Tomato	1	9.249 (0.592)	0.014 (0.014)	0.001 (0)	0.088 (0.187)	0.008 (0.011)	0.011 (0.033)	0.005 (0)	2.253 (1.805)	14.877 (22.386)	0.833 (0.408)	0.018 (0.035)	0.022 (0.046)	0.035 (0.102)	0.044 (0.132)	0.059 (0.162)	0.129 (0.056)
	2	24.544 (2.504)	0.050 (0.050)	0.001 (0)	0.131 (0.240)	0.009 (0.006)	0.013 (0.025)	0.006 (0.001)	2.852 (0.848)	13.540 (9.779)	1.079 (0.576)	0.020 (0.018)	0.024 (0.026)	0.037 (0.054)	0.054 (0.106)	0.084 (0.169)	0.091 (0.027)
	3	43.651 (4.829)	0.360 (0.360)	0.001 (0)	0.500 (0.534)	0.033 (0.006)	0.036 (0.007)	0.021 (0.006)	2.880 (4.579)	69.698 (311.362)	1.078 (0.128)	0.076 (0.012)	0.090 (0.014)	0.113 (0.020)	0.137 (0.031)	0.192 (0.064)	0.089 (0.030)
	4	53.727 (5.782)	1.140 (1.140)	0.001 (0)	0.611 (0.374)	0.060 (0.030)	0.061 (0.038)	0.038 (0.011)	1.718 (0.403)	4.884 (3.454)	0.974 (0.126)	0.136 (0.080)	0.160 (0.097)	0.198 (0.120)	0.232 (0.137)	0.306 (0.159)	0.108 (0.033)
	5	46.474 (3.091)	0.827 (0.827)	0.001 (0)	0.305 (0.134)	0.017 (0.005)	0.021 (0.010)	0.010 (0.002)	4.061 (1.885)	37.573 (41.772)	1.220 (0.213)	0.039 (0.013)	0.048 (0.018)	0.065 (0.029)	0.084 (0.045)	0.139 (0.075)	0.060 (0.023)

Appendix C

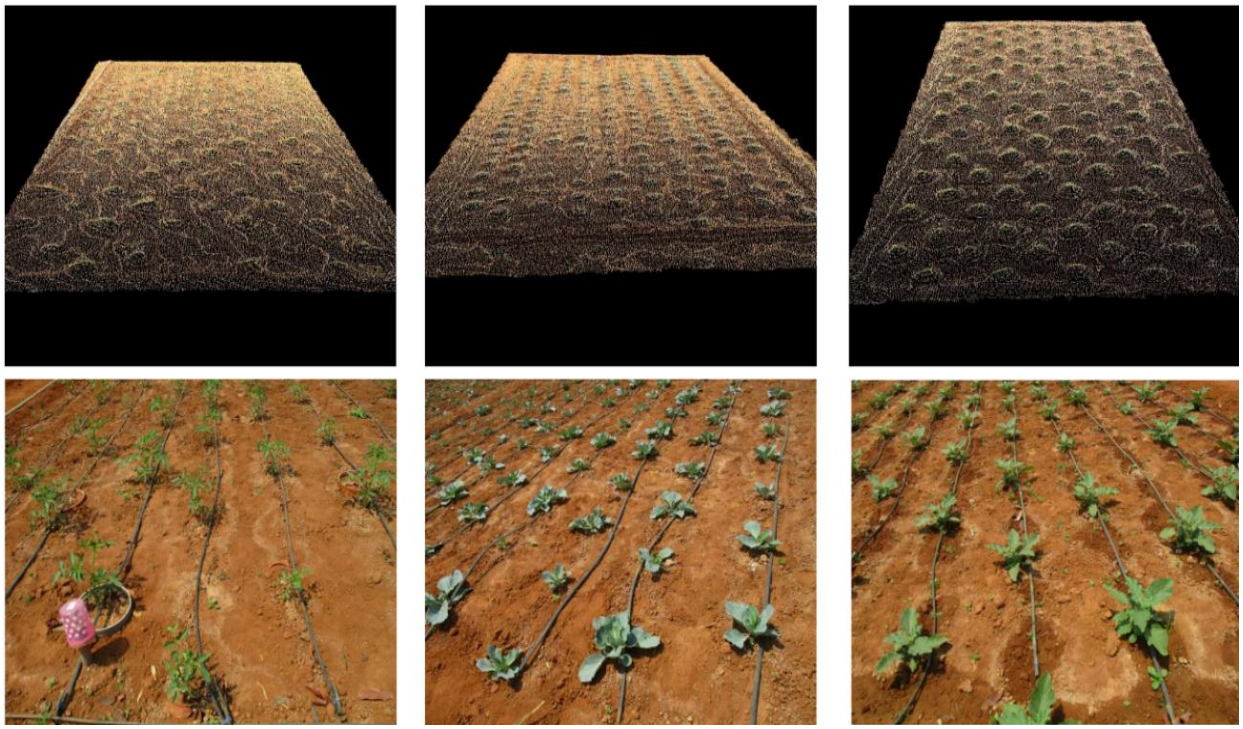


Figure A1. Exemplary pictures of three crops used in the study. Top: Point clouds. Bottom: Photographs. From left to right: Tomato, cabbage, and eggplant. All images and point clouds were collected at the second sampling date.

References

1. Malambo, L.; Popescu, S.C.; Murray, S.C.; Putman, E.; Pugh, N.A.; Horne, D.W.; Richardson, G.; Sheridan, R.; Rooney, W.L.; Avant, R.; et al. Multitemporal field-based plant height estimation using 3D point clouds generated from small unmanned aerial systems high-resolution imagery. *Int. J. Appl. Earth Observ. Geoinf.* **2018**, *64*, 31–42.

2. Zhang, C.; Kovacs, J.M. The application of small unmanned aerial systems for precision agriculture: A review. *Precis. Agric.* **2012**, *13*, 693–712.

3. Selsam, P.; Schaeper, W.; Brinkmann, K.; Buerkert, A. Acquisition and automated rectification of high-resolution RGB and near-IR aerial photographs to estimate plant biomass and surface topography in arid agro-ecosystems. *Exp. Agric.* **2017**, *53*, 144–157.

4. Yu, N.; Li, L.; Schmitz, N.; Tian, L.F.; Greenberg, J.A.; Diers, B.W. Development of methods to improve soybean yield estimation and predict plant maturity with an unmanned aerial vehicle based platform. *Remote Sens. Environ.* **2016**, *187*, 91–101.

5. Park, S.; Ryu, D.; Fuentes, S.; Chung, H.; Hernández-Montes, E.; O'Connell, M. Adaptive Estimation of Crop Water Stress in Nectarine and Peach Orchards Using High-Resolution Imagery from an Unmanned Aerial Vehicle (UAV). *Remote Sens.* **2017**, *9*, 828.

6. Johnson, C.K.; Mortensen, D.A.; Wienhold, B.J.; Shanahan, J.F.; Doran, J.W. Site-specific management zones based on soil electrical conductivity in a semiarid cropping system. *Agronomy J.* **2003**, *95*, 303–315.

7. Lati, R.N.; Filin, S.; Eizenberg, H. Estimating plant growth parameters using an energy minimization-based stereovision model. *Comput. Electron. Agric.* **2013**, *98*, 260–271.

8. Madec, S.; Baret, F.; de Solan, B.; Thomas, S.; Dutartre, D.; Jezequel, S.; Hemmerlé, M.; Colombeau, G.; Comar, A. High-throughput phenotyping of plant height: Comparing unmanned aerial vehicles and ground LiDAR estimates. *Front. Plant Sci.* **2017**, *8*, 2002.

9. Hoffmeister, D.; Waldhoff, G.; Korres, W.; Curdt, C.; Bareth, G. Crop height variability detection in a single field by multi-temporal terrestrial laser scanning. *Precis. Agric.* **2016**, *17*, 296–312.

10. Tilly, N.; Hoffmeister, D.; Cao, Q.; Huang, S.; Lenz-Wiedemann, V.; Miao, Y.; Bareth, G. Multitemporal crop surface models: Accurate plant height measurement and biomass estimation with terrestrial laser scanning in paddy rice. *J. Appl. Remote Sens.* **2014**, *8*, 083671.

11. Fricke, T.; Richter, F.; Wachendorf, M. Assessment of forage mass from grassland swards by height measurement using an ultrasonic sensor. *Comput. Electron. Agric.* **2011**, *79*, 142–152.

12. Bendig, J.; Bolten, A.; Bennertz, S.; Broscheit, J.; Eichfuss, S.; Bareth, G. Estimating biomass of barley using crop surface models (CSMs) derived from UAV-based RGB imaging. *Remote Sens.* **2014**, *6*, 10395–10412.

13. Leberl, F.; Irschara, A.; Pock, T.; Meixner, P.; Gruber, M.; Scholz, S.; Wiechert, A. Point clouds. *Photogramm. Eng. Remote Sens.* **2010**, *76*, 1123–1134.

14. Tumbo, S.D.; Salyani, M.; Whitney, J.D.; Wheaton, T.A.; Miller, W.M. Investigation of laser and ultrasonic ranging sensors for measurements of citrus canopy volume. *Appl. Eng. Agric.* **2002**, *18*, 367–372.

15. Li, W.; Niu, Z.; Chen, H.; Li, D. Characterizing canopy structural complexity for the estimation of maize LAI based on ALS data and UAV stereo images. *Int. J. Remote Sens.* **2017**, *38*, 2106–2116.

16. Weiss, M.; Baret, F. Using 3D point clouds derived from UAV RGB imagery to describe vineyard 3D macro-structure. *Remote Sens.* **2017**, *9*, 111.

17. Schirrmann, M.; Giebel, A.; Gleiniger, F.; Pflanz, M.; Lentschke, J.; Dammer, K.-H. Monitoring agronomic parameters of winter wheat crops with low-cost UAV imagery. *Remote Sens.* **2016**, *8*, 706.

18. Parkes, S.D.; McCabe, M.F.; Al-Mashhawari, S.K.; Rosas, J. Reproducibility of crop surface maps extracted from Unmanned Aerial Vehicle (UAV) derived Digital Surface Maps. In *Remote Sensing for Agriculture, Ecosystems, and Hydrology XVIII*; Neale, C.M.U., Maltese, A., Eds.; SPIE: Bellingham, WA, USA, 2016.

19. Prasad, J.V.N.S.; Rao, C.S.; Srinivas, K.; Jyothi, C.N.; Venkateswarlu, B.; Ramachandrappa, B.K.; Dhanapal, G.N.; Ravichandra, K.; Mishra, P.K. Effect of ten years of reduced tillage

and recycling of organic matter on crop yields, soil organic carbon and its fractions in Alfisols of semiarid tropics of southern India. *Soil Tillage Res.* **2016**, *156*, 131–139.

20. Lowe, D.G. Distinctive image features from scale-invariant key points. *Int. J. Comput. Vis.* **2004**, *60*, 91–110.
21. Snavely, N.; Seitz, S.M.; Szeliski, R. Modeling the world from internet photo collections. *Int. J. Comput. Vis.* **2008**, *80*, 189–210.
22. Westoby, M.J.; Brasington, J.; Glasser, N.F.; Hambrey, M.J.; Reynolds, J.M. 'Structure-from-Motion' photogrammetry: A low-cost, effective tool for geoscience applications. *Geomorphology* **2012**, *179*, 300–314.
23. Mesas-Carrascosa, F.-J.; Torres-Sánchez, J.; Clavero-Rumbao, I.; García-Ferrer, A.; Peña, J.-M.; Borra-Serrano, I.; López-Granados, F. Assessing optimal flight parameters for generating accurate multispectral orthomosaicks by uav to support site-specific crop management. *Remote Sens.* **2015**, *7*, 12793–12814.
24. Röder, M.; Hill, S.; Latifi, H. *Best Practice Tutorial: Technical Handling of the UAV "DJI Phantom 3 Professional" and Processing of the Acquired Data*; Dept. of Remote Sensing, University of Würzburg: Würzburg, Germany, 2017.
25. Tucker, C.J. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sens. Environ.* **1979**, *8*, 127–150.
26. Motohka, T.; Nasahara, K.N.; Oguma, H.; Tsuchida, S. Applicability of green-red vegetation index for remote sensing of vegetation phenology. *Remote Sens.* **2010**, *2*, 2369–2387.
27. Silva, C.A.; Hudak, A.T.; Klauberg, C.; Vierling, L.A.; Gonzalez-Benecke, C.; de Padua Chaves Carvalho, S.; Rodriguez, L.C.E.; Cardil, A. Combined effect of pulse density and grid cell size on predicting and mapping aboveground carbon in fast-growing Eucalyptus forest plantation using airborne LiDAR data. *Carbon Balance Manag.* **2017**, *12*, 13.
28. Næsset, E.; Økland, T. Estimating tree height and tree crown properties using airborne scanning laser in a boreal nature reserve. *Remote Sens. Environ.* **2002**, *79*, 105–115.

29. Roussel, J.-R.; Auty, D. lidR: Airborne LiDAR Data Manipulation and Visualization for Forestry Applications. Available online: https://CRAN.R-project.org/package=lidR (accessed on 18.05.2018).

30. R Core Team. *R*: *A Language and Environment for Statistical Computing*; R Core Team: Vienna, Austria, 2016. Available online: https://www.R-project.org/ (accessed on 18.05.2018).

31. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32.

32. Cortes, C.; Vapnik, V. Support-vector networks. *Mach. Learn.* **1995**, *20*, 273–297.

33. Liaw, A.; Wiener, M. Classification and Regression by random Forest. *R News* **2002**, *2*, 18–22.

34. Meyer, D.; Dimitriadou, E.; Hornik, K.; Weingessel, A.; Leisch, F. e1071: Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071), TU Wien, 2015. Available online: https://CRAN.R-project.org/package=e1071 (accessed on 18.05.2018).

35. Li, W.; Niu, Z.; Chen, H.; Li, D.; Wu, M.; Zhao, W. Remote estimation of canopy height and aboveground biomass of maize using high-resolution stereo images from a low-cost unmanned aerial vehicle system. *Ecol. Indic.* **2016**, *67*, 637–648.

36. Garcia-Gutierrez, J.; Martínez-Álvarez, F.; Troncoso, A.; Riquelme, J.C. A comparative study of machine learning regression methods on LiDAR data: A case study. In *International Joint Conference SOCO'13-CISIS'13-ICEUTE'13*; Herrero, Á., Baruque, B., Klett, F., Abraham, A., Snášel, V., de Carvalho, A.C.P.L.F., Bringas, P.G., Zelinka, I., Quintián, H., Corchado, E., Eds.; Springer International Publishing: Cham, Switzerland, 2014; pp. 249–258.

37. Horning, N. Random forests: An algorithm for image classification and generation of continuous fields data sets. In *Proceeding of the "International Conference on Geoinformatics for Spatial Infrastructure Development in Earth and Allied Sciences"*; Hanoi University of Mining and Geology: Hanoi, Vietnam, 2010.

38. Niederheiser, R.; Rutzinger, M.; Bremer, M.; Wichmann, V. Dense image matching of terrestrial imagery for deriving high-resolution topographic properties of vegetation locations in alpine terrain. *Int. J. Appl. Earth Observ. Geoinf.* **2018**, *66*, 146–158.

39. Cunliffe, A.M.; Brazier, R.E.; Anderson, K. Ultra-fine grain landscape-scale quantification of dryland vegetation structure with drone-acquired structure-from-motion photogrammetry. *Remote Sens. Environ.* **2016**, *183*, 129–143.

40. Iqbal, F.; Lucieer, A.; Barry, K.; Wells, R. Poppy Crop Height and Capsule Volume Estimation from a Single UAS Flight. *Remote Sens.* **2017**, *9*, 647.

41. Bendig, J.; Yu, K.; Aasen, H.; Bolten, A.; Bennertz, S.; Broscheit, J.; Gnyp, M.L.; Bareth, G. Combining UAV-based plant height from crop surface models, visible, and near infrared vegetation indices for biomass monitoring in barley. *Int. J. Appl. Earth Observ. Geoinf.* **2015**, *39*, 79–87.

42. Moeckel, T.; Safari, H.; Reddersen, B.; Fricke, T.; Wachendorf, M. Fusion of ultrasonic and spectral sensor data for improving the estimation of biomass in grasslands with heterogeneous sward structure. *Remote Sens.* **2017**, *9*, 98.

43. Safari, H.; Fricke, T.; Reddersen, B.; Moeckel, T.; Wachendorf, M. Comparing mobile and static assessment of biomass in heterogeneous grassland with a multi-sensor system. *J. Sens. Sens. Syst.* **2016**, *5*, 301–312.

Chapter 3

Multi-temporal monsoon crop biomass estimation using hyperspectral imaging

Abstract

Hyperspectral remote sensing is considered to be an effective tool in crop monitoring and estimation of biomass. Many of the previous approaches are from single year or single date measurements, even though the complete crop growth with multiple years would be required for an appropriate estimation of biomass. The aim of this study was to estimate the fresh matter biomass (FMB) by terrestrial hyperspectral imaging of the three crops lablab, maize and finger millet under different levels of nitrogen (N) fertiliser and water supply. Further, the importance of the different spectral regions for the estimation of FMB was assessed. The study was conducted in two experimental layouts (rainfed (R) and irrigated (I)) at the University of Agricultural Sciences, GKVK campus, Bengaluru, India. Spectral images and the FMB were collected over three years (2016–2018) during the growing season of the crops. Random forest regression method was applied to build FMB models. R^2 validation (R^2_{val}) and relative root mean square error prediction (rRMSEP) was used to evaluate the FMB models. The Generalised model (combination of R and I data) performed better for lablab (R^2_{val} = 0.53, rRMSEP = 13.9%), maize (R^2_{val} = 0.53, rRMSEP = 18.7%) and finger millet (R^2_{val} = 0.46, rRMSEP = 18%) than the separate FMB models for R and I. In the best derived model, the most important variables contributing to the estimation of biomass were in the wavelength ranges of 546–910 nm (lablab), 750–794 nm (maize) and 686–814 nm (finger millet). The deviation of predicted and measured FMB did not differ much among the different levels of N and water supply. However, there was a trend of overestimation of FMB at the initial stage and underestimation at the later stages of crop growth.

Introduction

The majority of India's population (60%) depends on the agricultural sector for their livelihood [1]. Agriculture depends mainly on monsoon rainfall, surface water and ground water irrigation. Since the variability of monsoon rainfall is high, it forces the farmers to adapt their irrigated areas to local water availability [2]. Irrigated crop production is a major contributor to the green revolution, which has enabled the country to be self-sufficient [3], accompanied by fertiliser application,

chemical weed and pest control. Timely fertiliser application with water supply is essential for a successful crop. Spectral data from Remote Sensing (RS) have been studied for many years for an assessment of nutrient and water variability for yield optimisation [4–6].

RS can be an effective tool in monitoring crop production [7–9] and estimating yield [10,11]. Early estimation of yield may allow better planning and forecasting the market prices and support food security based on the regional, national and global demand and supply. RS allows collecting information about crop production using non-destructive methods [12] on a large scale for many fields at the same time. Hyperspectral (HS) RS provides continuous narrow spectral data from 400 to 2500 nm and have been proved to capture the variations in spectral response of the crop for the detection of nitrogen (N) content [13,14], biomass [15] and water stress [6,16]. Development of HS sensors and their application in estimating crop biomass from multi-year data [17] has gained increasing attention in recent years. Multi-temporal images provide more information on vegetation phenology under wet and dry conditions than a single image [18]. Many studies related to multi-temporal hyperspectral imaging have been published on crops such as rice (*Oryza sativa* L.) [19], wheat (*Triticum aestivum* L.) [20,21], and maize (*Zea mays* L.) [10]. Besides maize, lablab (*Lablab purpureus* L.), and finger millet (*Eleusine coracana* L.) are primary crops in the semi-arid region of South India. The state of Karnataka generates the major share of lablab (90%) [22] and finger millet (62.01%) [23] production of India. However, well-defined multi-year studies on the estimation of biomass for maize, lablab and finger millet using multi-temporal hyperspectral data under varying nitrogen (N) fertiliser and water supply levels are lacking.

The aim of this study was to assess the potential of terrestrial hyperspectral imaging for the estimation of monsoon crop biomass based on data from three years (2016–2018). The specific objectives of the study were: (1) to develop statistical models to predict the fresh matter biomass (FMB) of the three crops: lablab, maize and finger millet; (2) to assess the effect of different levels of N and water supply on the predicted FMB value for a wide range of crop phenology over the complete growing period; and (3) to evaluate the importance of spectral regions in the resulting models and understand the causal relationships of the model.

Methods

The study was conducted during 2016–2018 at GKVK campus of University of Agricultural Sciences, Bengaluru (UASB) located in the eastern dry zone of Karnataka state, India

(12°58′20.79″N, 77°34′50.31″E, 920 m.a.s.l) (Figure 7a). The soils of the study area are formed by Kandic Paleustalfs and Dystric Nitisols as dominant soil types and the climate is described as a tropical savanna climate with a rainy season from June to October. Mean annual temperature is 29.2 °C with an average precipitation of 923 mm [24]. The total rainfall and mean temperature data of the monsoon cropping season varied from 2016 to 2018 (Table 4) [25].

Table 4. Total rainfall and mean temperature data of the cropping seasons in a multi-factorial field experiment at University of Agricultural Sciences, GKVK Campus, Bengaluru, India.

Years	**2016**	**2017**	**2018**
Rainfall (mm)	403.4	763.2	264.8
Temperature (°C)	23.51	23.48	23.28

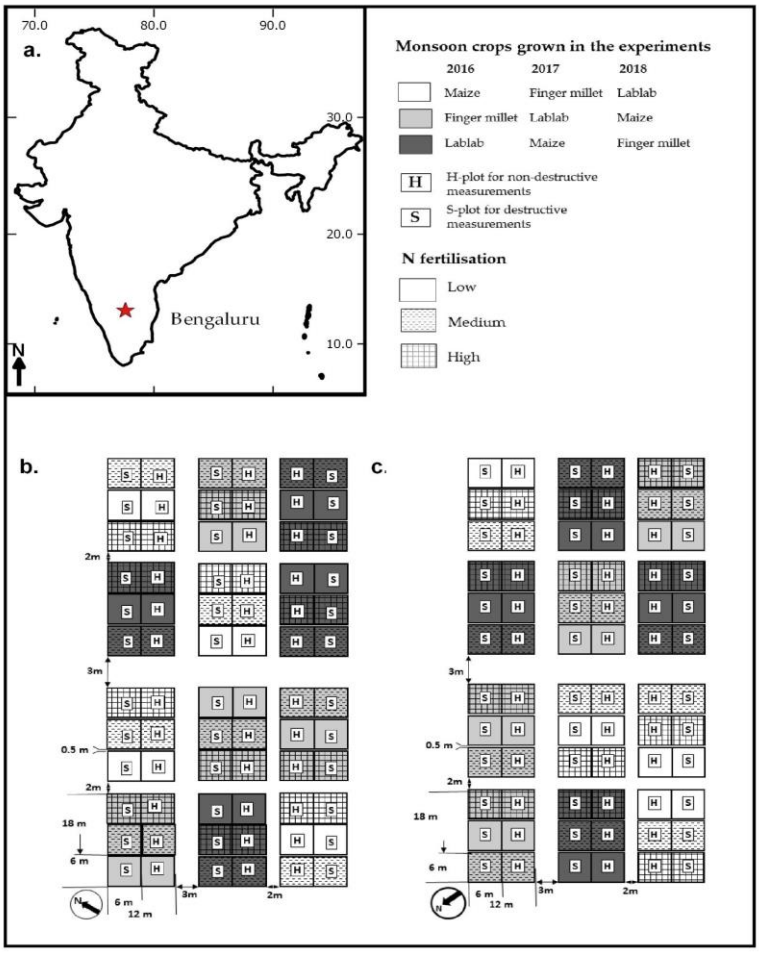


Figure 7. (a) Location of Bengaluru within India; (b) design of rainfed experimental layout; (c) design of irrigated experimental layout (adjusted from [11]).

Experimental Site

Two experiments were established with different irrigation regimes: rainfed (R) and drip irrigated (I). Each year has two cropping seasons, namely rainy/monsoon season (July–November) and dry season (February–May). Even in the monsoon season, drip irrigation systems are common, as the southwest monsoons are getting increasingly unreliable and timely irrigation enhances crop productivity. In the rainy season, lablab (cultivars: HA 4 and HA 3), maize (cultivars: Nithyashree and NAH 1137) and finger millet (cultivars: GPU-28, MR-6 and ML-365) were cultivated in both experiments (Table A1) [26–30]. Fertiliser application was done by broadcasting at three levels of N fertilizer to all crops. At high levels, the complete 100% application of N (recommended dosage) was applied and a reduced amount was applied at medium and low levels, which varied across years (Table A2). To lablab, the complete N dosage was applied at the time of sowing and for maize and finger millet it was split into two halves, i.e. at the time of sowing and four weeks after sowing (top dressing) with the objective of supplying nitrogen to growing plants in the readily available form and avoid leaching losses by heavy rainfalls, which frequently occur after sowing. Phosphorous (P) and potassium (K) were applied completely to all crops at the time of sowing (Table A2).

Each block of a particular crop had three experimental plots (6 m x 12 m) with three N levels (low, medium and high) arranged in a randomised block design (Figure 7). In this split plot experiment, each plot was subdivided into two sub plots (6 m each), one was used for destructive biomass sampling (S) and the other one was used for non-destructive spectral measurements (H). In total, 36 plots (3 crops x 4 blocks x 3 fertiliser levels) were used for the spectral and biomass sampling (Figure 7).

Spectral Data Measurements

Three hyperspectral images were taken in each H subplot using the full-frame hyperspectral camera UHD 185-Firefly [31] mounted on a terrestrial tripod. The distance between the camera and the plant canopy height was 1.5 m throughout the growth of the crop to cover the same area of approximately 1 m² in all images. The camera measured the spectral range from 450 to 998 nm, of which the wavelength between 470 and 950 nm was further analysed, as the signal-to-noise ratio was too low for the wavebands from 450 to 470 and 950 to 998 nm. The spectral range was divided into 121 bands with a band width of 4 nm. The focal length of the camera was 12.1 mm with the

image size of 50 x 50 pixels covering the area of 1 m x 1 m at the applied height (2 cm spatial resolution). Prior to the measurement, the camera was calibrated using a dark (cap covering the lens) and white reflectance plane (95% reflectance Zenith Lite) [32] to calculate reflectance directly from the measured radiance. Although light conditions varied throughout the three years due to different illumination angles, the goal was to keep it as constant as possible. During the calibration in the field, the integration time was automatically obtained. The spectral reflectance of each pixel was normalised by dividing with the maximum reflectance value of the same pixel to reduce temporal variation and random noise [33].

Each image contained non-vegetation elements such as soil, drip irrigation pipes and shadows. To reduce the effect of these elements, a two-step procedure was applied (Figure 8). First, the Normalised Difference Vegetation Index (NDVI) [8] was calculated as the difference between reflectance in the red (620–750 nm) and near-infrared (NIR) (750–1400 nm), divided by the sum of reflectance in the red and NIR spectral range. Second, a two-class k-means clustering algorithm was applied to separate vegetation and non-vegetation using the NDVI values. The two class centroids were identified based on NDVI values in such a way that the distance between the centroids were minimised. Finally, only the pixels classified as vegetation were used to calculate the average spectral reflectance for each image. The three images collected in each plot were averaged resulting in one spectral reflectance curve per H subplot and sampling date.

Biomass Sampling

The fresh matter weight of 2–4 plants were determined in the field and extrapolated to 1 m². The maximum number of samples collected for a particular crop type in a particular growing season was 60 (5 sampling dates × 3 fertiliser treatments × 4 replicates), but the number of sampling varied among the three years (Table 5).

Table 5. Total number of samples from multi-factorial field experiments rainfed (R) and irrigated (I) at University of Agricultural Sciences, GKVK Campus, Bengaluru, India.

Experiments.	**Number of samples in R.**			**Number of samples in I**		
Year	2016	2017	2018	2016	2017	2018
Lablab	60	60	12	52	60	12
Maize	48	48	12	60	48	12
Finger millet	60	36	12	52	36	12
Total annual crop-wise	168	144	36	164	144	36
Total experiment-wise	348			344		
Grand total	692					

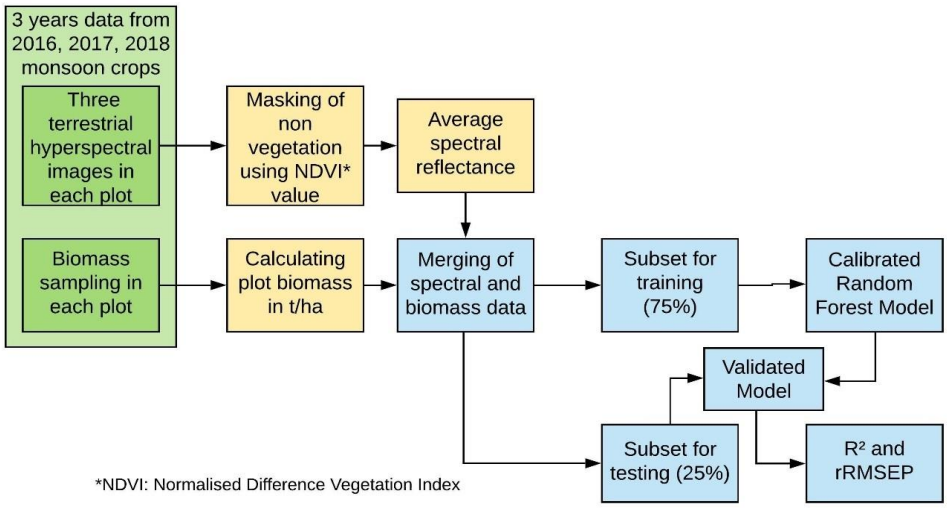


Figure 8. Workflow showing the data collection (green), data preparation (yellow) and data analysis (blue).

Sampling Dates

In 2016 (Y1), the sampling was done on five sampling dates (Y1S1–Y1S5) under both irrigation regimes (R and I) for the three crops. Rainfed maize at the final sampling date in 2016 (Y1S5) was not sampled due to technical difficulties with the sensor. In 2017 (Y2), there were five sampling dates (Y2S1–Y2S5) for lablab, four for maize (Y2S1–Y2S4) and three for finger millet (Y2S1–Y2S3) in both irrigation regimes (R and I). In Y2, irrigated maize top-dressing fertilisation was mixed up for a few low, medium and high plots and hence the plots I07, I08, I09, I13, I14, and I15 (Figure 7c) were eliminated from analysis for Sampling dates 2–4 (Y2S2–Y2S4). In 2018 (Y3), there was one sampling date (Y3S1) for all three crops. Assessment of phenological stages of the crop was carried out by recording the morphological characteristics of the plants according to Biologische Bundesanstalt, Bundessortenamt und Chemische Industrie (BBCH) [34]. In total, there were 11 sampling dates (BBCH 1–8) for 2016–2018 (Table A3).

Statistical Analysis

To predict the fresh matter biomass (FMB) from reflectance data, machine learning random forest regression (RFR) in caret package [35] was used [36]. As the original dataset was skewed towards one side, the FMB measured in the S subplots were cube root transformed to assure normal distribution of the dataset although RFR does not require normal distribution of the FMB dataset. RFR is a regression tree technique, which builds multiple decision trees and ensembles them for an accurate prediction [36]. It is less sensitive to overfitting as the subsets are drawn randomly each

time. The regression trees has the ability to deal with complex relationships between variables for large datasets [37]. Each crop was modelled separately for the R and I experiment based on datasets of three years (6 models). Further, the datasets from both irrigation regimes on each crop were combined to check the robustness of the model independent of water supply in one Generalised model. To eliminate the bias involved in splitting the data into training and testing sets, 100 different random subsets (75% for training and 25% for testing) were generated based on the sampling dates for each crop separately. Using these random subsets, 100 RFR models were calibrated and validated to predict the FMB for each crop from reflectance data. All machine learning methods have specific configuration parameters called tune parameters or hyperparameters, which optimise the performance of the predictive modelling algorithm [38]. For RFR model, two tune parameters need to be determined, i.e. number of trees and mtry. The number of tree parameters was always kept to a default value of 500 and the mtry parameter value was tuned using the repeated cross-validation (five-fold, three repeats) procedure. The mtry parameter value was set between 1 and 15 and the optimum mtry parameter for each model was identified. The model estimation accuracy of FMB was evaluated using R^2 validation (R^2val) (Equation (1)) [39], the root mean square error of prediction (RMSEP) (Equation (2)), and the relative root mean square error prediction (rRMSEP) (Equation (3)).

$$R^2_{val} = \left[1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y}_i)^2}\right]$$

$$RMSEP = \sqrt{\frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{n}}$$

$$rel.RMSEP = \frac{RMSEP}{\max(y_i) - \min(y_i)}$$

where y_i is the measured fresh matter biomass, $\hat{y}_i$ is the predicted fresh matter biomass, $\bar{y}_i$ is the average measured fresh matter biomass, and n is the number of samples.

To determine the important wavelengths in the prediction of FMB, the best model was identified out of 100 FMB models on each crop based on the lowest RMSE value. From the best model, the wavelengths contributing above 75% in the prediction of FMB were identified. The normalised deviation between the predicted and observed FMB values were calculated and differences in the deviation based on N levels, sampling dates and water supply were examined (Equation (4)).

$$Normalised\ deviation = \frac{Predicted\ FMB\ -\ Observed\ FMB}{Predicted\ FMB\ +\ Observed\ FMB}$$

Results

In the rainfed experiment, the range of FMB (S sub-plot) over the three years 2016–2018 was 0.16–14.6 t/ha for lablab, 0.76–67.71 t/ha for maize and 0.89–59.39 t/ha for finger millet (Figure 9). Similarly, for irrigated experiment, it was 0.22–44.33 t/ha for lablab, 2.28–79.38 t/ha for maize and 0.91–69.63 t/ha for finger millet. Crop growth continuously increased until S3 or S4 and started to decrease at later stages in all crops and along the three years. The FMB was higher in I than R experiment except for finger millet at Y1S1 and Y2S2.

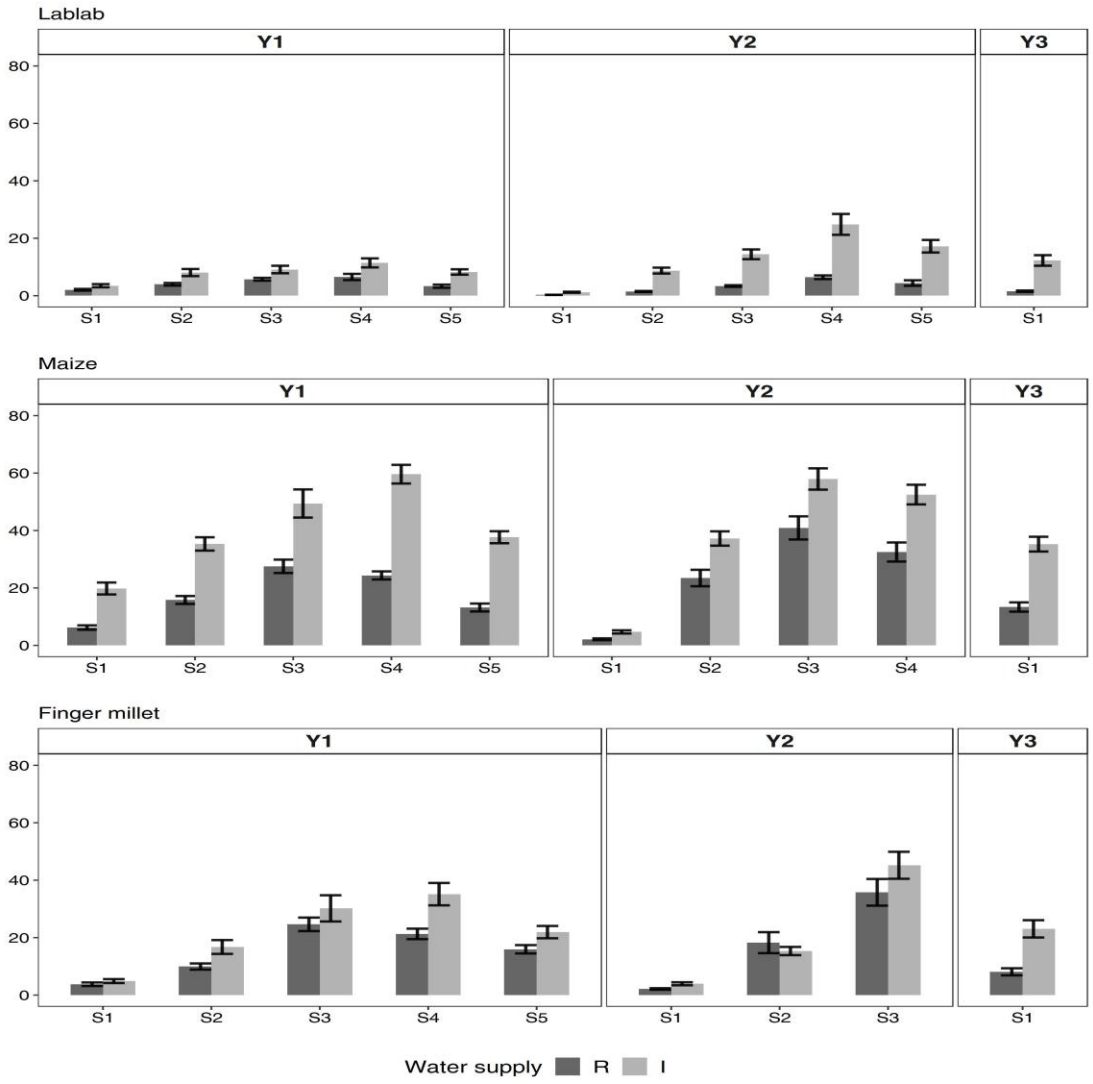


Figure 9. Fresh matter biomass (FMB) in the rainfed (R) and irrigated (I) multi-factorial field experiments at University of Agricultural Sciences, GKVK Campus, Bengaluru, India from 2016–2018 (Y1–Y3). The diagrams show average values over three levels of N fertiliser application.

To gain an impression of the spectral variation for each crop, minimum, average and maximum spectral reflectance from the images of rainfed and irrigated experiments were determined for the three crops lablab, maize and finger millet during the three monsoon seasons (Figure 10).

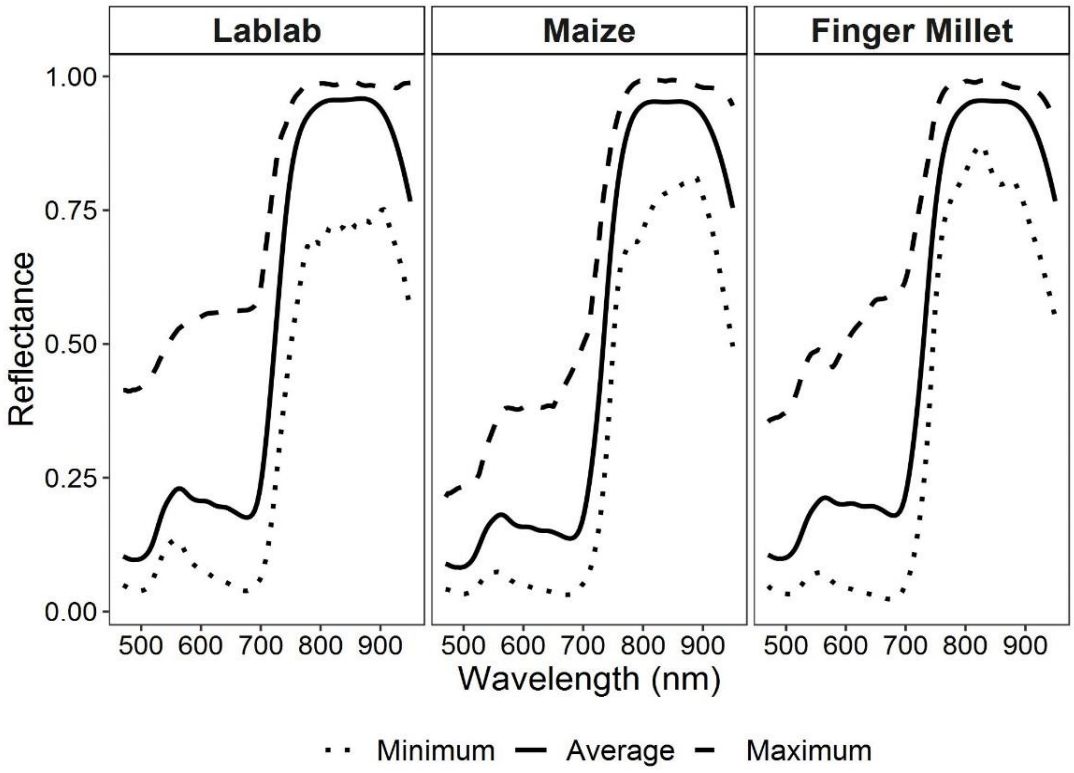


Figure 10. Minimum, average and maximum spectral reflectance curves of lablab, maize and finger millet for three levels of N and two levels of water supply during the three monsoon seasons.

Crop Specific FMB Models

To develop a prediction model for FMB, which is valid for varying conditions, individual FMB models were developed for two irrigation regimes (i.e., R and I) and a combination of the datasets of both irrigation regimes (i.e., Generalised model). The prediction accuracy of the FMB models varied between the crops and depended on the dataset (R, I, or Generalised) used for model development (Figure 11). The lowest rRMSEP value nearing to zero was considered as a better model. Building the RFR models separately for both R and I treatments, the lowest median rRMSEP for lablab was found in I with 17.9% (R^2_{val} = 0.34) and for maize and finger millet in R experiment with 18.5% (R^2_{val} = 0.60) and 19.8% (R^2_{val} = 0.46), respectively. With the combined dataset, the rRMSEP for lablab was 13.9% (R^2_{val} = 0.53), for finger millet was 18% (R^2_{val} = 0.46) and for maize

was 18.7% (R^2_{val} = 0.53). Overall, compared to the experiment-wise modelling approach, model accuracy (in terms of rRMSEP) was higher for all crops when models were built with data from both water supply levels.

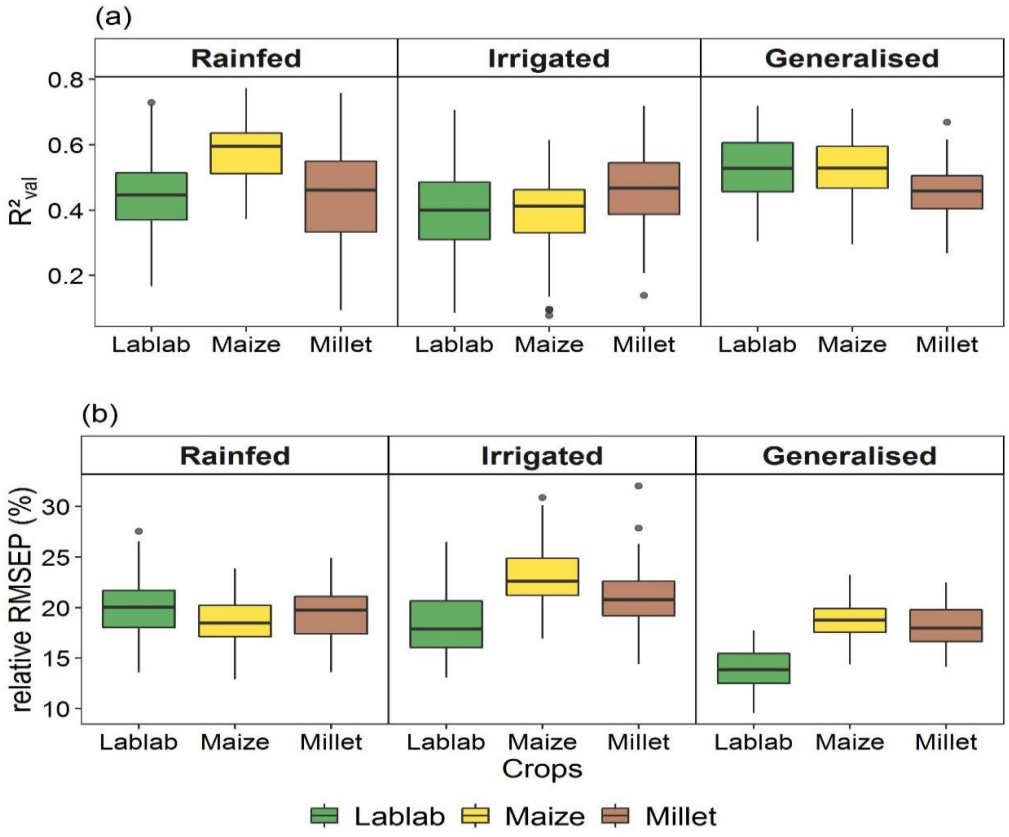


Figure 11. Prediction accuracy measured as R^2val (a) and rRMSEP (b) values of the models (Rainfed, Irrigated and Generalised) for fresh matter biomass of lablab, maize and finger millet. Models were built on data from three different years, three levels of N and two levels of water supply (i.e., rainfed and irrigated).

In RFR modelling, the mtry parameter indicates the number of input variables randomly chosen at each node. Optimum mtry values (best tune values) were found to be 13, 7 and 7 for lablab; 8, 12 and 13 for maize; and 8, 2 and 7 for finger millet, respectively, for the Rainfed, Irrigated and Generalised models.

The randomised models were based on stratified (according to sampling date and fertilisation rate) randomly selected samples for the calibration and validation dataset (Figure 12). Having considered such random effects in RFR modelling, it becomes obvious that predictions show a substantial underestimation with increasing FMB values (Figure 12).

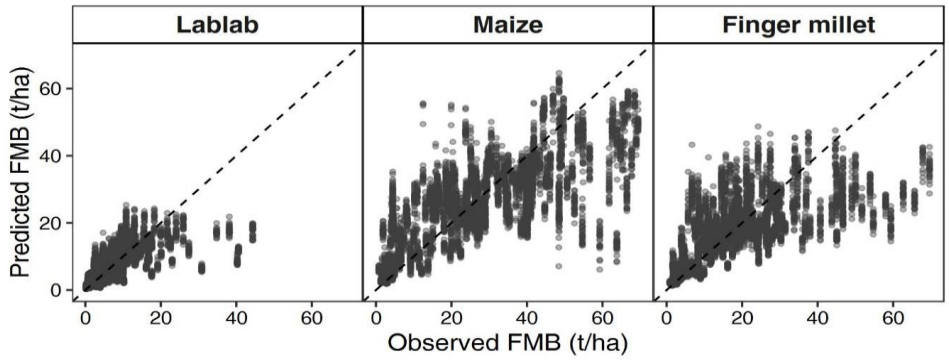


Figure 12. Plot of fit of the Generalised models for fresh matter biomass (FMB) of lablab, maize and finger millet. Each plot shows predictions from 100 RFR models with randomly selected calibration and validation data. Models were built on data from three different years andthree levels of N in multi-factorial field experiments rainfed (R) and irrigated (I) at University of Agricultural Sciences, GKVK Campus, Bengaluru, India.

Performance of the Generalised Models Considering N Application Rates, Sampling Dates and Water Supply

The normalised deviation of predicted and measured biomass was used to check if the prediction accuracy of Generalised models varied among the three levels of N application (Figure 13). Overall, only minor deviations were found among low, medium and high levels of N supply for all crops between 2016 and 2018.

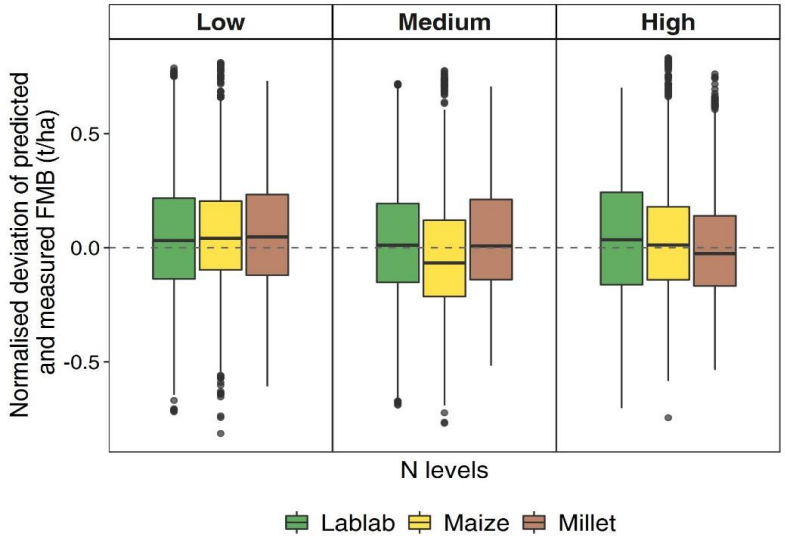


Figure 13. Normalised deviation between predicted and measured biomass for lablab, maize, and finger millet at three levels of N application (low, medium and high). Predictions were based on the Generalised model. Values were averaged over 11 sampling dates (2016–2018) from multi-factorial field experiments rainfed (R) and irrigated (I) at University of Agricultural Sciences, GKVK Campus, Bengaluru, India.

Prediction accuracy of Generalised models varied strongly among the sampling dates (Figure 14). While in Y1, normalised deviation for lablab showed an irregular pattern, such as an overestimation (Y1S1 and Y1S3), underestimation (Y1S2) and good concordance (Y1S4 and Y1S5). A decreasing trend of deviation was observed with increasing crop maturity in Y2. With maize, there was a general decline across sampling dates both in Y1 and Y2. With finger millet, there was overestimation for the early sampling dates (Y1S1–Y1S2 and Y2S1–Y2S2) followed by decreasing underestimation for the later sampling dates in 2016 (Y1S3–Y1S5). Following this deviation, it can be concluded that crop phenology influenced model performance with a tendency towards overestimation of biomass at early stages and an underestimation at later stages of crop growth.

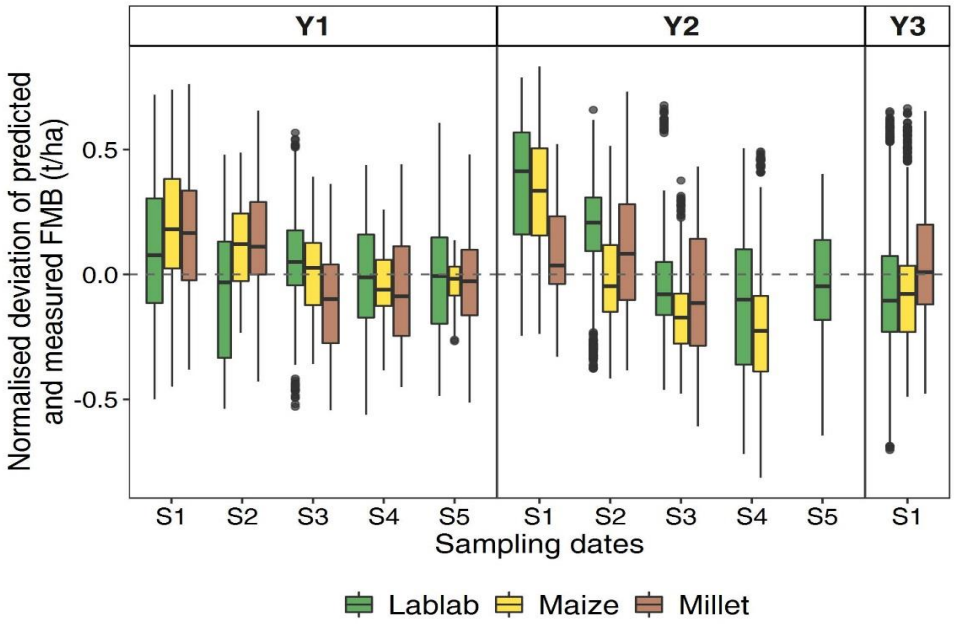


Figure 14. Normalised deviation between predicted and measured fresh matter biomass (FMB) for lablab, maize and finger millet at each sampling date (S1–S5) over three years (Y1–Y3). Predictions were based on the Generalised model. Values were averaged over 11 sampling dates (2016–2018) from multi-factorial field experiments rainfed (R) and irrigated (I) at University of Agricultural Sciences, GKVK Campus, Bengaluru, India.

No systematic over- or underestimation was found for biomass prediction of the three crops at the two levels of water supply (Figure 15). Hence, model prediction was rather robust with slightly larger deviations for lablab at both water supply levels as compared to the other crops.

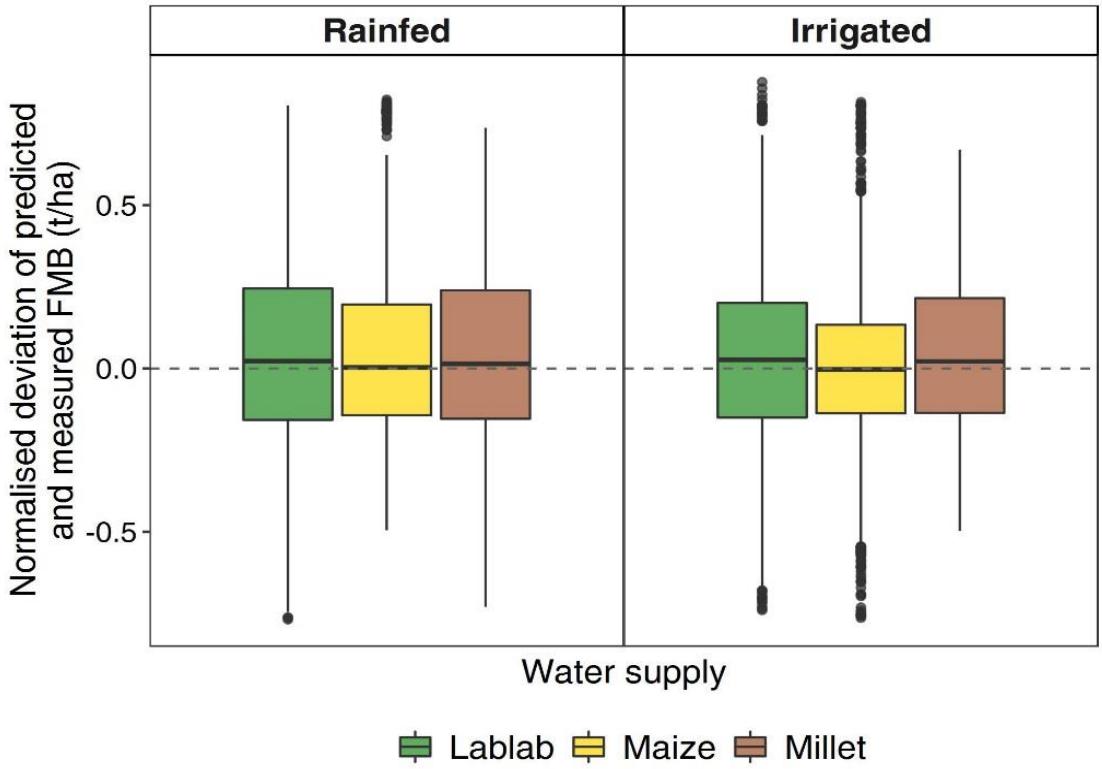


Figure 15. Normalised deviation between predicted and measured biomass for lablab, maize and finger millet at two levels of water supply (rainfed and irrigated). Predictions were based on the Generalised model. Values were averaged over 11 sampling dates (2016–2018) from multi-factorial field experiments rainfed (R) and irrigated (I) at University of Agricultural Sciences, GKVK Campus, Bengaluru, India.

Importance of Wavelengths

Wavelengths of the crop reflectance helped in differentiating and identifying the crop traits based on their spectral region. The best model was identified out of 100 Generalised models on each crop based on the lowest RMSE value. From the best model, the wavelengths contributing above 75% in the prediction of FMB were identified (Figure 16). For lablab, a multitude of spectral bands from the green, red and near infrared (NIR) region (546–910 nm) contributed significantly to the estimation of biomass. Contrastingly, for maize, only wavelengths in the NIR region (750–794 nm) and for finger millet in both the red and NIR region (686, 694, and 774–814 nm) were important.

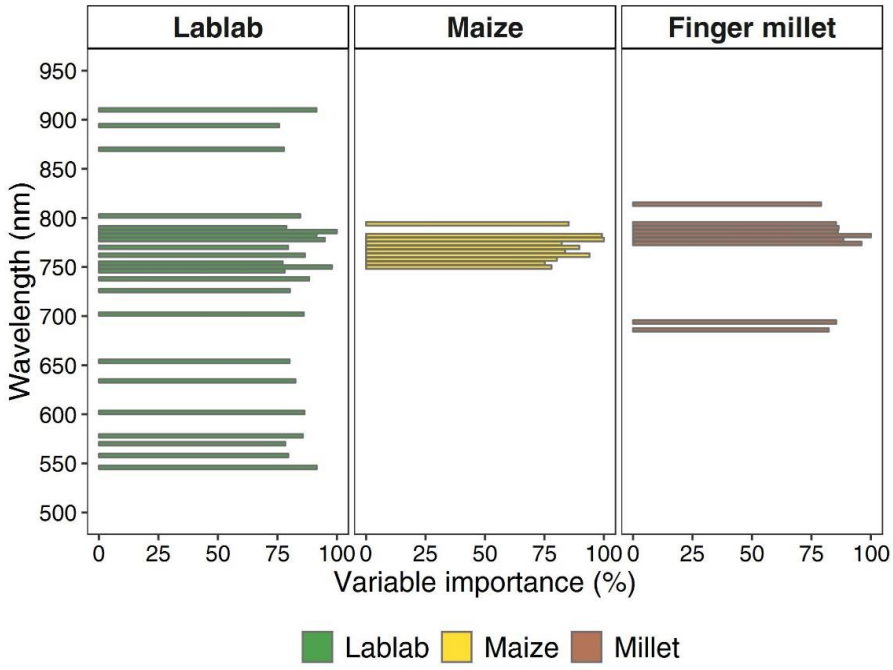


Figure 16. Important wavelengths (score above 75) in the Generalised models for fresh matter biomass of lablab, maize and finger millet. Models were built on data from three different years (2016-2018) andthree levels of N from multi-factorial field experiments rainfed (R) and irrigated (I) at University of Agricultural Sciences, GKVK Campus, Bengaluru, India.

Discussion

The aim of the study was to estimate the monsoon crop biomass for three crops (lablab, maize and finger millet) based on terrestrial hyperspectral imaging during crop growth season across three years (2016–2018). With a high number of samplings during three consecutive monsoon seasons, a wide range of phenological stages of crops could be covered. This is important considering the validity range of prediction models, since the harvest time of crops varies considerably in agricultural practice, for example due to nutrient and water availability and moisture content of grains. Thus, by our deliberate multi-temporal approach, the validity range of Generalised models was significantly broadened, which was further enhanced by the integration of crop measurements under a wide range of N fertiliser and water supply.

The FMB models were developed based on the predicted FMB values and tested with the observed FMB values for validation. Overall, the results indicate that the Generalised models had higher estimation accuracy (with rRMSEP ranging from 13.9% to 18.7%) for all the three crops, as compared to the rainfed and irrigated models. One reason may be that, with the combination of data

from two experiments, representing severe water limitation (Rainfed experiment) and optimum water supply (Irrigated experiment), the range of crop productivity became much broader, which eventually may have increased the robustness of regression models.

Similar prediction errors were found in a previous study for maize biomass by RGB images (relative error 16.66%, $R^2 = 0.78$) [40]. In contrast to our study, their models included canopy height parameter additional to RGB information, which shows the promising potential of structural data calculated with photogrammetric methods particularly when they are combined with data from other sensor types [11,41,42]. Although spectral information was limited to the Red Edge Modified Ratio Index (REMRI), the combination of spectral data with LiDAR-derived metrics produced only a slightly smaller error in the estimation of maize biomass [10] as compared to our study. However, as the sampling was done at only one date of one single year and because no defined N and water supply was applied, the transferability of such modelling approaches beyond the study area may be limited.

Although lablab is an important legume in the food and cattle production system in India, this plant has not been subjected to any remote sensing assessment this far. The fact that it was the least productive crop in both experiments across all years, strongly reduced the range of FMB values for model calibration. However, the highest prediction errors obtained were between those of the more productive crops maize and finger millet. Similarly, finger millet is a rarely researched crop in terms of remote sensing. In a single-year satellite-based study with pearl millet, which exhibits a similar growth pattern as finger millet, Lambert et al. [43] found a strong relationship between Sentinel-2 based LAI data and crop biomass ($R^2 = 0.84$), which is much higher than in our study ($R^2 = 0.46$). Although neither sensors and platforms, nor the range of crop phenology and management were comparable, this study highlights the scope of well-informed satellite-based hyperspectral imagery, and proximal imagery may make important contributions to such developments, e.g., by the provision of crop-specific spectral libraries as a source of reference spectra that can aid the interpretation of hyperspectral and multispectral image [44].

Although we observed quite some deviation between predicted and observed FMB, the median was close to zero at all levels of N and water supply, when the Generalised models were used for all the three crops. This proves the robustness of models, which allow biomass prediction irrespective of varying nitrogen and water management practices. However, the pronounced pattern of

deviations along the sampling dates in Y1 and Y2 points at the limitations of models, which are solely built on spectral information. Although soil-containing pixels were masked out of the images prior to model calibration, a substantial overestimation of biomass at the initial sampling dates in the growing season occurred, while biomass was frequently underestimated at later sampling dates. The overestimation of biomass may be caused by weeds at the initial sampling dates as the effect of weeds could not be controlled in the estimation of biomass. Further, the prediction error for crops increased in the order lablab (13.9%), finger millet (18%) and maize (18.7%), which clearly shows that spectral information captured at the top canopy layer is increasingly less representative of the biomass at lower layers of the canopy. This effect is also addressed as the "saturation constraint" and was regularly found in previous studies (e.g., [45–47]) particularly when vegetation indices, such as the Normalised Differential Vegetation Index (NDVI), were used. Obviously, this problem cannot be circumvented by the use of individual spectral wavelengths instead of vegetation indices, but stresses the vital necessity to develop multi-sensor approaches, in which each sensor's shortcomings are compensated by other sensors [10,48,49].

As a common trait for all three crops, wavelengths in the red-edge area were of utmost importance for the estimation of crop biomass. The Generalised model for lablab further comprised several wavelengths in the green, red and NIR region, indicating a larger number of variables in these models. Similar important bands were found by Manjunath et al. [50] in the discrimination of chickpea, pea and lentils. While in maize the most important variables were found in the red-edge region, the model of finger millet also contained wavelengths in the red region as important variables. For lablab, several bands were identified in the visible part of spectrum (450–750 nm) to be important for biomass prediction. These bands are known to be affected by plant pigments, especially by chlorophyll [51]. The ability of lablab to fix atmospheric nitrogen may have resulted in longer greenness of the leaves over the growing period, which leads to a higher reflectance at the green peak (~550 nm) and a higher absorbance in red (~650 nm). In general, the identified spectral bands confirm accepted knowledge about biomass-reflectance relationships [52].

Potential and Limitations

Although Generalised models performed better at various management practices, the application of terrestrial hyperspectral measurements is still time consuming and cannot be applied on larger scales. Contrarily, drone techniques carry great potential to collect hyperspectral imagery in a

comparable spatial resolution for larger areas. Another limitation is the dependence of the relative prediction error of the models from the growth development of the crops, which may have been enhanced by the change of crop varieties across years.

Conclusions

The results show that random forest regression modelling based on multi-temporal hyperspectral imagery allows the prediction of fresh matter biomass of three major food and feed crops in the monsoon season of southern India. The results of this study showed that Generalised models, which were built on crop data from both rainfed and irrigated conditions, are more robust than water management specific models. For all Generalised crop models, deviations between predicted and observed values were independent of N fertiliser and water supply, indicating a wide validity range of the models. However, an overestimation of crop biomass was detected at initial growth stages of crops along with an underestimation at the later stages of the crop growth, which was particularly pronounced with the more productive crops maize and finger millet. While wavelengths in the red edge region were important variables in all three Generalised crop models, several others in the visible and near infrared region were important in models for lablab and finger millet. The results of this study suggest that, for the tested monsoon crops at advanced maturity, even hyperspectral information is not sufficient for an accurate biomass prediction. Data fusion from a combination of sensors may improve the prediction performance, as complementary sensors can compensate for their respective deficiencies.

Appendix A

Table A1. Varietal description of the crops lablab, maize and finger millet grown in three years (2016-2018) from multi-factorial field experiments rainfed (R) and irrigated (I) at University of Agricultural Sciences, GKVK Campus, Bengaluru, India.

Crops	Years grown	Varieties	Sowing	Duration (days)	Yield (t/ha)		Salient features
					Rainfed	Irrigated	
Lablab	2016 and 2017	HA 4 (HA 3 x Magadi local)	Can be grown throughout the year as they are photo insensitive	100–105	Dry Seeds: 1–1.2, Green pods: 4.5–5		Pods are constricted with characteristic odour (Sogadu) in all the three cropping seasons
	2018	HA 3 (HA 1 x US 67-31)		95–100	Dry Seeds: 0.8–0.9 Green pods: 4.5–5		Flat pods with no odour (Sogadu)
Maize	2016 and 2017	Nithyashree (SKV-50 x NA1-105)	Can be grown throughout the year	110–120	Grain: 7.41–7.90 Straw: 19.77	Grain: 7.90–8.40 Straw: 29.65	Tolerant to downy mildew, leaf blight and stem borer
	2018	NAH 1137 (Hema)		110–120	Grain: 7.90–8.40 Straw: 19.77	Grain: 8.89–9.39 Straw: 29.65	
Finger millet	2016	GPU-28 (Indaf 5 x (Indaf 9 x IE 1012))	July–August	110–115	Average Grain: 3.5–4		Medium tall plants, semi compact ears with tip incurved fingers. Highly resistant to finger and neck blast
	2017	MR-6 (African white x RoH 2)	June–July	120–125	Average Grain: 3–3.5		100–110 cm tall plants, open ears with tip incurved fingers
	2018	ML-365 (IE 1012 x Indaf 5)	June–August (Kharif monsoon) January–February (Rabi dry)	110–115	Average Grain: 5–5.5		Medium height, semi compact ears with tip incurved fingers. Resistant to neck blast and tolerant to drought

Table A2. Nitrogen (N), phosphorous (P_2O_5) and potassium (K_2O) application rates to lablab, maize and finger millet crops in multi-factorial field experiments rainfed (R) and irrigated (I) at University of Agricultural Sciences, GKVK Campus, Bengaluru, India during 2016–2018.

Mineral fertilisation	Lablab						Maize						Finger millet					
	2016		2017		2018		2016		2017		2018		2016		2017		2018	
	R	I	R	I	R	I	R	I	R	I	R	I	R	I	R	I	R	I
N (kg ha^{-1}) §	25	25	25	25	25	25	100	150	100	150	150	150	50	100	50	100	50	50
P_2O_5 (kg ha^{-1})	10	10	50	50	10	10	50	75	50	75	50	75	40	50	40	50	40	50
K_2O (kg ha^{-1})	10	10	25	25	10	10	37.5	50	25	40	37.5	50	37.5	50	37.5	50	37.5	50

§ Recommended application rate (referred to as high in Figure 1); averaged across years, medium application rates were 53.3%, 56.2% and 58.3% from the high level for lablab, maize and finger millet, respectively; low application rates in 2016 were 40.0%, 41.7% and 50.0% from the high level for lablab, maize and finger millet, respectively, and zero application was done in 2017 and 2018.

Table A3. Phenological stages of lablab, maize and finger millet at the sampling dates in multi-factorial field experiments rainfed (R) and irrigated (I) at University of Agricultural Sciences, GKVK Campus, Bengaluru, India from 2016 to 2018 (BBCH scale).

Rainfed experiment									
Sampling dates	**Lablab (BBCH/DAS*)**			**Maize (BBCH/DAS*)**			**Finger millet (BBCH/DAS*)**		
	2016	**2017**	**2018**	**2016**	**2017**	**2018**	**2016**	**2017**	**2018**
1	2/40	2/23		3/42	1/25		2/44	2/30	
2	5/53	2/38		5/61	3/45		3/65	3/52	
3	6/63	5/47		7/81	6/67	7/79	5/94	5/81	5/79
4	7/73	6/69	7/78	8/98	7/108		7/109		
5	8/89	8/89					8/126		
Irrigated experiment									
1	2/41	2/24		3/43	1/27		2/45	2/32	
2	6/56	5/39		5/62	3/46		3/66	3/53	
3	7/64	6/48		7/88	6/68	7/87	5/96	5/82	5/87
4	7/74	7/72		7/100	7/110		7/110		
5	8/97	8/90	8/83	8/128			8/135		

* DAS, Days after Sowing.

BBCH:
1: Leaf development
2: Formation of side shoots/tillering
3: Stem elongation
5: Inflorescence emergence
6: Flowering
7: Development of fruit
8: Ripening

References

1. Arjun, K.M. Indian Agriculture-Status, Importance and Role in Indian Economy. *Int. J. Agric. Food Sci. Technol.* **2013**, *4*, 343–346.
2. Ferrant, S.; Selles, A.; Le Page, M.; Herrault, P.A.; Pelletier, C.; Al-Bitar, A.; Mermoz, S.; Gascoin, S.; Bouvet, A.; Saqalli, M.; et al. Detection of irrigated crops from Sentinel-1 and Sentinel-2 data to estimate seasonal groundwater use in South India. *Remote Sens.* **2017**, *9*, 1119.
3. Thenkabail, P.; Dheeravath, V.; Biradar, C.; Gangalakunta, O.R.P.; Noojipady, P.; Gurappa, C.; Velpuri, M.; Gumma, M.; Li, Y. Irrigated area maps and statistics of India using remote sensing and national statistics. *Remote Sens.* **2009**, *1*, 50–67.
4. Cohen, Y.; Alchanatis, V.; Zusman, Y.; Dar, Z.; Bonfil, D.J.; Karnieli, A.; Zilberman, A.; Moulin, A.; Ostrovsky, V.; Levi, A.; et al. Leaf nitrogen estimation in potato based on spectral data and on simulated bands of the VENμS satellite. *Precis. Agric.* **2010**, *11*, 520–537.
5. Samborski, S.M.; Tremblay, N.; Fallon, E. Strategies to make use of plant sensors-based diagnostic information for nitrogen recommendations. *Agron. J.* **2009**, *101*, 800–816.
6. Zhang, F.; Zhou, G. Estimation of vegetation water content using hyperspectral vegetation indices: A comparison of crop water indicators in response to water stress treatments for summer maize. *BMC Ecol.* **2019**, *19*, 18.
7. Aasen, H.; Burkart, A.; Bolten, A.; Bareth, G. Generating 3D hyperspectral information with lightweight UAV snapshot cameras for vegetation monitoring: From camera calibration to quality assurance. *ISPRS J. Photogramm. Remote Sens.* **2015**, *108*, 245–259.
8. Rouse, J.W.; Haas, R.H.; Schell, J.A.; Deering, D.W. Monitoring vegetation systems in the great plains with ERTS. *NASA Spec. Publ.* **1974**, *1*, 309–317.
9. Warren, G.; Metternicht, G. Agricultural applications of high-resolution digital multispectral imagery: Evaluating within-field spatial variability of Canola (Brassica napus) in Western Australia. *Photogramm. Eng. Remote Sens.* **2005**, *71*, 595–602.
10. Wang, C.; Nie, S.; Xi, X.; Luo, S.; Sun, X. Estimating the biomass of maize with hyperspectral and LiDAR data. *Remote Sens.* **2017**, *9*, 11.

11. Moeckel, T.; Dayananda, S.; Nidamanuri, R.; Nautiyal, S.; Hanumaiah, N.; Buerkert, A.; Wachendorf, M. Estimation of vegetable crop parameter by multi-temporal UAV-borne images. *Remote Sens.* **2018**, *10*, 805.

12. Burkart, A.; Aasen, H.; Alonso, L.; Menz, G.; Bareth, G.; Rascher, U. Angular dependency of hyperspectral measurements over wheat characterized by a novel UAV based goniometer. *Remote Sens.* **2015**, *7*, 725–746.

13. Yu, K.; Li, F.; Gnyp, M.L.; Miao, Y.; Bareth, G.; Chen, X. Remotely detecting canopy nitrogen concentration and uptake of paddy rice in the Northeast China Plain. *ISPRS J. Photogramm. Remote Sens.* **2013**, *78*, 102–115.

14. Vigneau, N.; Ecarnot, M.; Rabatel, G.; Roumet, P. Potential of field hyperspectral imaging as a non destructive method to assess leaf nitrogen content in wheat. *F. Crop. Res.* **2011**, *122*, 25–31.

15. Yue, J.; Feng, H.; Jin, X.; Yuan, H.; Li, Z.; Zhou, C.; Yang, G.; Tian, Q. A comparison of crop parameters estimation using images from UAV-mounted snapshot hyperspectral sensor and high-definition digital camera. *Remote Sens.* **2018**, *10*, 1138.

16. Krishna, G.; Sahoo, R.N.; Singh, P.; Bajpai, V.; Patra, H.; Kumar, S.; Dandapani, R.; Gupta, V.K.; Viswanathan, C.; Ahmad, T.; et al. Comparison of various modelling approaches for water deficit stress monitoring in rice crop through hyperspectral remote sensing. *Agric. Water Manag.* **2019**, *213*, 231–244.

17. Koppe, W.; Gnyp, M.L.; Hennig, S.D.; Li, F.; Miao, Y.; Chen, X.; Jia, L.; Bareth, G. Multi-temporal hyperspectral and radar remote sensing for estimating winter wheat biomass in the North China Plain. *Photogramm. - Fernerkundung - Geoinf.* **2012**, *3*, 281–298.

18. Knight, J.F.; Lunetta, R.S.; Ediriwickrema, J.; Khorram, S. Regional scale land cover characterization using MODIS-NDVI 250 m multi-temporal imagery: A phenology-based approach. *GIScience Remote Sens.* **2006**, *43*, 1–23.

19. Aasen, H.; Gnyp, M.L.; Miao, Y.; Bareth, G. Automated Hyperspectral Vegetation Index Retrieval from Multiple Correlation Matrices with HyperCor. *Photogramm. Eng. Remote Sens.* **2014**, *80*, 785–795.

20. Honkavaara, E.; Saari, H.; Kaivosoja, J.; Pölönen, I.; Hakala, T.; Litkey, P.; Mäkynen, J.; Pesonen, L. Processing and assessment of spectrometric, stereoscopic imagery collected using a lightweight UAV spectral camera for precision agriculture. *Remote Sens.* **2013**, *5*, 5006–5039.

21. Yue, J.; Yang, G.; Li, C.; Li, Z.; Wang, Y.; Feng, H.; Xu, B. Estimation of winter wheat above-ground biomass using unmanned aerial vehicle-based snapshot hyperspectral sensor and crop height improved models. *Remote Sens.* **2017**, *9*.

22. Byre Gowda, M. DOLICHOS BEAN - Lablab purpureus (L.) Sweet. Available online: http://www.lablablab.org/html/general-information.html (accessed on 15 May 2019).

23. Selected state-wise area, production and productivity of ragi in India (2016-17). Available online: https://www.indiastat.com/table/agriculture-data/2/ragi-finger-millet/17200/1131134/data.aspx (accessed on 16 May 2019).

24. Prasad, J.V.N.S.; Rao, C.S.; Srinivas, K.; Jyothi, C.N.; Venkateswarlu, B.; Ramachandrappa, B.K.; Dhanapal, G.N.; Ravichandra, K.; Mishra, P.K. Effect of ten years of reduced tillage and recycling of organic matter on crop yields, soil organic carbon and its fractions in alfisols of semi arid tropics of Southern India. *Soil Tillage Res.* **2016**, *156*, 131–139.

25. University of Agricultural Sciences Bangalore Agrometeorology Available online: https://www.uasbangalore.edu.in/index.php/research/agromoterology (accessed on 21 July 2019).

26. Byre Gowda, M. HA-4, a new variety of Lablab purpureus introduced for cultivation in Karnataka. In Proceedings of the Annual Plant Breeder's Conference; University of Agricultural Sciences, Bangalore, India, 23-25 October 2007.

27. Byre Gowda, M. Genetic enhancement of Dolichos bean through integration of conventional breeding & molecular approaches and farmers participatory plant breeding. 2011. Available online: https://www.kirkhousetrust.org/docs/reports/dolichos_report_april_2011.pdf (accessed on 22 June 2019).

28. Hybrid maize varietal description. Available online: http://e-krishiuasb.karnataka.gov.in/ItemDetails.aspx?depID=1&subDepID=1&cropID=4# (accessed on 14 June 2019).

29. Kaul, J.; Dass, S.; Manivannan, A.; Singode, A.; Sekhar, J.C.; Chikkappa, G.K.; Parkash, O. Maize hybrid and composite varieties released in India. *Tech. Bull.* 2010, 80.

30. Project Coordinating Unit Compendium of released varieties in small millets. 2014. Available online: https://www.dhan.org/smallmillets/docs/report/Compendium_of_Released_Varieties_in_Small_millets.pdf (accessed on 22 June 2019).

31. Hyperspectral imaging - real snapshot technology - Cubert GmbH. Available online: https://cubert-gmbh.com/ (accessed on 3 April 2019).

32. SphereOptics Zenith Lite™ Ultralight Targets Available online: http://sphereoptics.de/en/product/zenith-lite-ultralight-targets/ (accessed on 22 July 2019).

33. Cao, F.; Yang, Z.; Ren, J.; Jiang, M.; Ling, W.-K. Does normalization methods play a role for hyperspectral image classification? *Comput.* **2017,** *arXiv:1710,1–6*.

34. Zadoks, J.C.; Chang, T.T.; Konzak, C.F. A decimal code for the growth stages of cereals. *Weed Res.* **1974**, *14*, 415–421.

35. Max, K.; Wing, J.; Weston, S.; Williams, A.; Keefer, C.; Engelhardt, A.; Cooper, T.; Mayer, Z.; Kenkel, B.; the R Core Team; et al. caret: Classification and regression training 2018. Available online: https://cran.r-project.org/web/packages/caret/index.html (accessed on 22 June 2019).

36. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32.

37. Ismail, R.; Mutanga, O.; Kumar, L. Modeling the potential distribution of pine forests susceptible to Sirex Noctilio infestations in Mpumalanga, South Africa. *Trans. GIS* **2010**, *14*, 709–726.

38. Raschka, S. Model evaluation, model selection, and algorithm selection in machine learning. *Comput. Res. Repos.* **2018**, *arXiv:1811*, 1–49.

39. Kvalseth, T.O. Cautionary note about R^2. *Am. Stat.* **1985**, *39*, 279–285.

40. Li, W.; Niu, Z.; Chen, H.; Li, D.; Wu, M.; Zhao, W. Remote estimation of canopy height and aboveground biomass of maize using high-resolution stereo images from a low-cost unmanned aerial vehicle system. *Ecol. Indic.* **2016**, *67*, 637–648.

41. Bendig, J.; Bolten, A.; Bennertz, S.; Broscheit, J.; Eichfuss, S.; Bareth, G. Estimating biomass of barley using Crop Surface Models (CSMs) derived from UAV-based RGB imaging. *Remote Sens.* **2014**, *6*, 10395–10412.

42. Cunliffe, A.M.; Brazier, R.E.; Anderson, K. Ultra-fine grain landscape-scale quantification of dryland vegetation structure with drone-acquired structure-from-motion photogrammetry. *Remote Sens. Environ.* **2016**, *183*, 129–143.

43. Lambert, M.J.; Blaes, X.; Traore, P.S.; Defourny, P. Estimate yield at parcel level from S2 time serie in Sub-Saharan smallholder farming systems. In Proceedings of the 2017 9th international workshop, Analysis of Multitemporal Remote Sensing Images (MultiTemp), Brugge, Belgium, 14 November 2017.

44. Rama Rao, N.; Garg, P.K.; Ghosh, S.K. Development of an agricultural crops spectral library and classification of crops at cultivar level using hyperspectral data. *Precis. Agric.* **2007**, *8*, 173–185.

45. Pimstein, A.; Karnieli, A.; Bonfil, D.J. Wheat and maize monitoring based on ground spectral measurements and multivariate data analysis. *J. Appl. Remote Sens.* **2007**, *1*, 013530.

46. Freeman, K.W.; Girma, K.; Arnall, D.B.; Mullen, R.W.; Martin, K.L.; Teal, R.K.; Raun, W.R. By-plant prediction of corn forage biomass and nitrogen uptake at various growth stages using remote sensing and plant height. *Agron. J.* **2007**, *99*, 530–536.

47. Baez-Gonzalez, A.D.; Chen, P.; Tiscareno-Lopez, M.; Raghavan, S. Using satellite and field data with crop growth modeling to monitor and estimate corn yield in Mexico. *Crop Sci.* **2002**, *42*, 1943–1949.

48. Reddersen, B.; Fricke, T.; Wachendorf, M. A multi-sensor approach for predicting biomass of extensively managed grassland. *Comput. Electron. Agric.* **2014**, *109*, 247–260.

49. Kross, A.; McNairn, H.; Lapen, D.; Sunohara, M.; Champagne, C. Assessment of RapidEye vegetation indices for estimation of leaf area index and biomass in corn and soybean crops. *Int. J. Appl. Earth Obs. Geoinf.* **2015**, *34*, 235–248.

50. Manjunath, K.R.; Ray, S.S.; Panigrahy, S. Discrimination of spectrally-close crops using ground-based hyperspectral data. *J. Indian Soc. Remote Sens.* **2011**, *39*, 599–602.

51. Kokaly, R.F.; Asner, G.P.; Ollinger, S.V.; Martin, M.E.; Wessman, C.A. Characterizing canopy biochemistry from imaging spectroscopy and its application to ecosystem studies. *Remote Sens. Environ.* **2009**, *113*, S78–S91.

52. Ollinger, S.V. Sources of variability in canopy reflectance and the convergent properties of plants. *New Phytol.* **2011**, *189*, 375–394.

Chapter 4

Vegetable crop biomass estimation using hyperspectral in situ and RGB UAV data

Abstract

A sustainable agricultural intensification is necessary to cope with global population growth and increasing loss of agricultural land. Remote sensing (RS) has been an effective tool to monitor agricultural production systems and precision agriculture has been used for decades for field crops. Vegetable crops, on the other hand, have received less interest, although intensification in vegetable production requires particularly high levels of fertilizer and water.

The primary objective of this study was to test the predictive performance of two types of RS data – crop height information derived from point clouds, and reflectance information from hyperspectral imagery – to predict fresh matter yield (FMY) for three vegetable crops (eggplant, tomato, and cabbage). The study was conducted in an experimental layout at the University of Agricultural Sciences Bengaluru, India, using five sampling dates between May and June 2017. For the development of the biomass retrieval models, four machine learning (ML) methods were tested (partial least squares PLS, support vector machines SVM, random forest RFR, and gradient boosting trees GBT). Furthermore, the effect of the sampling date on the FMY prediction was examined.

The prediction accuracy for the three crops varied strongly depending on the ML method applied, and the RS dataset used. For all crops, RFR showed the best predictive performance, with an R^2_{val} of 0.97 for eggplant and tomato, and 0.93 for cabbage. The relative prediction error was below 10 % for the three crops. The RS dataset resulting in the most accurate prediction of biomass differed between the crops. For eggplant and cabbage, the spectral datasets showed the best prediction (R^2_{val} = 0.97 and 0.93, nRMSE = 0.07 and 0.09 for eggplant and cabbage). For tomato, the height dataset (R^2_{val} = 0.97, nRMSE = 0.06) clearly outperformed the spectral dataset alone and the fusion of both datasets. For eggplant and tomato, an increasing trend towards underestimation of the measured biomass is apparent for the later growth stages (sampling date 4 and 5). However, no significant ($p < 0.05$) effect of the growth stage on the biomass prediction was found.

Overall, the results of the study prove that an estimation of vegetable crop FMY using point clouds and hyperspectral imagery is successful throughout the growing season. Although the fusion of both datasets did not outperform single dataset use, the results indicate that, more than one sensory dataset should be collected to assure an optimal biomass prediction model.

Introduction

In the course of increasing urbanisation, sustainable resource use has become a pressing issue (Rockström et al., 2009). In developing countries, the population is projected to double by 2050 (UN, 2017), and the occupation of arable land by cities will be more than double. In combination with widespread malnutrition in such countries, especially in southern Asia, sustainable agriculture is necessary in order to increase food production (Tilman et al., 2011). For decades (e.g. Pettorelli et al., 2005; Tucker, 1979), remote sensing (RS) has been an effective tool to monitor agricultural production systems (Atzberger, 2013; Duncan et al., 2015). Particularly in countries with several cropping seasons per year, such as India, agricultural systems require multi-temporal RS data for a continuous evaluation of crop state and less ambiguous crop identification (Hannerz and Lotsch, 2008; Mondal et al., 2014).While crop identification can help to examine effects of e.g. urbanisation on the farmers' decisions to grow certain crops, an evaluation of the crop state can enable evaluation of the intensity (i.e. fertilization and irrigation) of the agricultural practice.

Intensive vegetable production, which is increasingly characterised by frequent nitrogen (N) fertilization (Thompson et al., 2018), requires rapid and frequent assessment of crop status (Padilla et al., 2018). While precision agriculture (PA) for field crops has been widely applied for two decades, it has a rather short history in horticulture (Zude-Sasse et al., 2016). The smaller field size and planting density as well as the complex plant architecture makes PA in horticulture complex (Zude-Sasse et al., 2016). Various sensors, from terrestrial (Nidamanuri and Zbell, 2011) and airborne (Moeckel et al., 2018) to satellite (Panigrahy and Sharma, 1997), have been used for investigating agro-environmental systems. Recently, unmanned aerial vehicles (UAV) have gained interest as platforms for collecting RS information from agricultural fields (Moeckel et al., 2018). UAVs allow collection of data from large areas in a highly repetitive manner, which makes them more independent from e.g. weather conditions than satellite or airborne platforms (Zhang and Kovacs, 2012). Another advantage of UAV platform is their flexibility to carry various sensor

systems, such as RGB cameras (Moeckel et al., 2018), multispectral sensors (Maimaitijiang et al., 2017), and hyperspectral sensors (Aasen and Bolten, 2018), for monitoring crop parameters.

The availability of different sensor systems enables sensor fusion, i.e. the combination of datasets from two or more sensors with different characteristics (Pohl and van Genderen, 1998). Sensor fusion may allow a more holistic interpretation of the relationship between RS information and crop parameters. For example, the fusion of spectral reflectance information with data from an ultra-sound system showed an increased prediction performance for barley (*Hordeum vulgare* L.) yield (Rischbeck et al., 2016). Other studies have shown an increased prediction performance of vegetation parameters, such as leaf area index (LAI) and biomass for various crops, e.g. maize (*Zea mays* L.) (Gao et al., 2013), winter wheat (*Triticum aestivum* L.) (Yue et al., 2018), and soybean (*Glycine max* L.) (Maimaitijiang et al., 2017), when spectral and structural information (based e.g. on RADAR, LiDAR, or photogrammetric methods) were combined.

With the development of new sensor systems, the amount of data generated has increased tremendously, creating a demand for statistical methods which can handle a large quantity and often redundancy of data. Machine learning (ML) methods represent a group of empirical statistical methods able to handle huge amounts of data and intercorrelation of variables. Additionally, ML methods have frequently been used for modelling crop biomass with RS data (e.g. Maimaitijiang et al., 2017; Moeckel et al., 2018; Rischbeck et al., 2016). Various ML methods have been introduced in the RS community in recent years (Wu et al., 2016), e.g. partial least squares (PLS) regression (Wold et al., 2008) and random forests (RFR) (Breiman, 2001). Although a large variety of ML approaches are available, none of them consistently outperforms the others, thus requiring that multiple methods be tested in order to find the optimal statistical model.

The primary objective of this study was to test the potential of two types of RS data – RGB (red-green-blue) data acquired by a UAV and terrestrial hyperspectral data – to predict fresh matter biomass of three vegetable crops (eggplant, tomato, and cabbage). For establishing the empirical relationship, four different ML approaches were tested (partial least squares, support vector machines, random forest, and gradient boosting trees regression). The contribution and potential of data from each sensor individually (crop height values from the RGB dataset and spectral information from the hyperspectral dataset) and from both sensors fused together were evaluated. We hypothesized that sensor fusion improves the predictive performance of the models for all crops.

Additionally, we expect that one ML approach alone will not deliver the optimal model all the time.

Methods

The data for this study was collected on the experimental farms of the University of Agricultural Sciences (UAS) in Bengaluru, India (12°58′20.79″N, 77°34′50.31″E, 920 asl) (Figure 17). The dominant soil types are Kandic Paleustalfs and Dystric Nitisols. The experimental site is characterised by a tropical climate with a distinct monsoon season from June to October. The mean annual temperature is 29.2 °C with an average precipitation of 923 mm (Prasad et al., 2016).

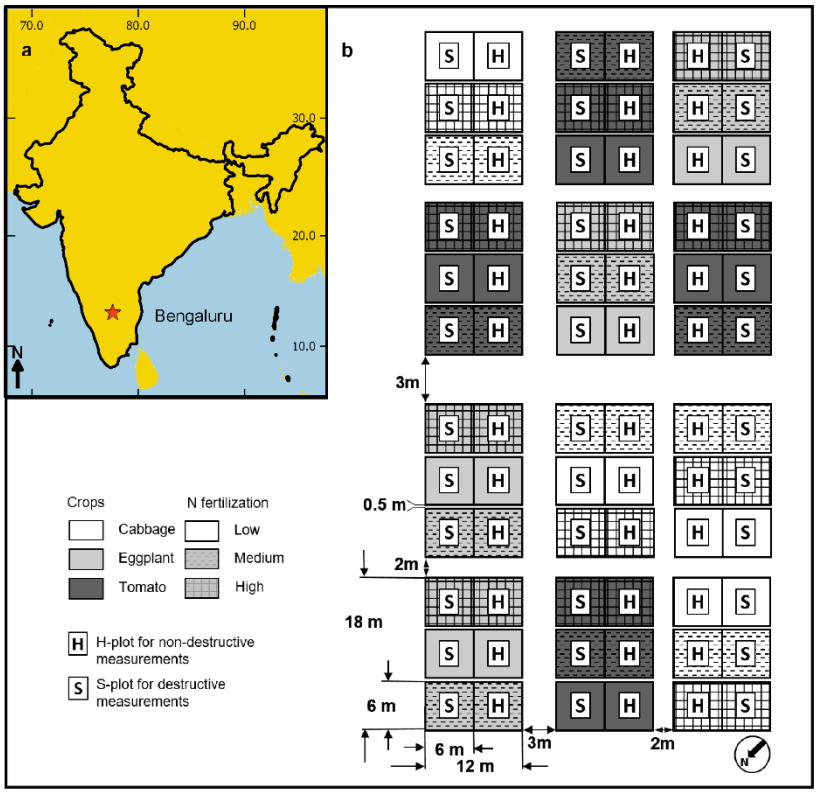


Figure 17: a) Position of Bengaluru within India. b) Design of the experimental layout (adjusted from Moeckel et al., 2018).

Field experiment

The field experiment was established in 2016 and had a distinct rainy and dry season crop rotation. In the dry season (January to May), tomato (*Solanum lycopersicum* L.; cultivar: NS-501), eggplant (*Solanum melongena* L.; cultivar: Ankur), and cabbage (*Brassica oleracea* L.; cultivar: Unnati)

were grown (Fig. 1B) with the use of a drip irrigation system. Twelve main plots (12 × 18 m) were arranged in a split-plot experiment. Each treatment was replicated four times (Figure 17B). Three subplots (6 × 12 m) with three levels of nitrogen (N) fertilizer were randomized within each main plot (Moeckel et al., 2018). As fertilizer was directly applied to individual plants, no nutrient deficiency was noticeable in any of the treatments. To allow for destructive measurements without disturbing neighbouring plants, each subplot was divided equally into a destructive sub-subplot for biomass sampling (S-plot; Figure 17B) and a non-destructive sub-subplot (H-plot) for spectral measurements. The S- and H-plots were 6 × 6 m each. The layout was thus comprised of a total of 36 subplots (3 crops × 3 fertilizer treatments × 4 replicates).

Plant sampling

The total biomass (vegetative parts as well as fruits at later growth stages) was measured at five sampling dates between March and May 2017. Two randomly selected plants were harvested (2 cm above the ground) in each S-plot (Figure 17b). At each sampling date, the biomass (t ha^{-1}) was measured in the field. The total number of samples collected for each crop type was 60 (5 sampling dates × 3 fertilizer treatments × 4 replicates).

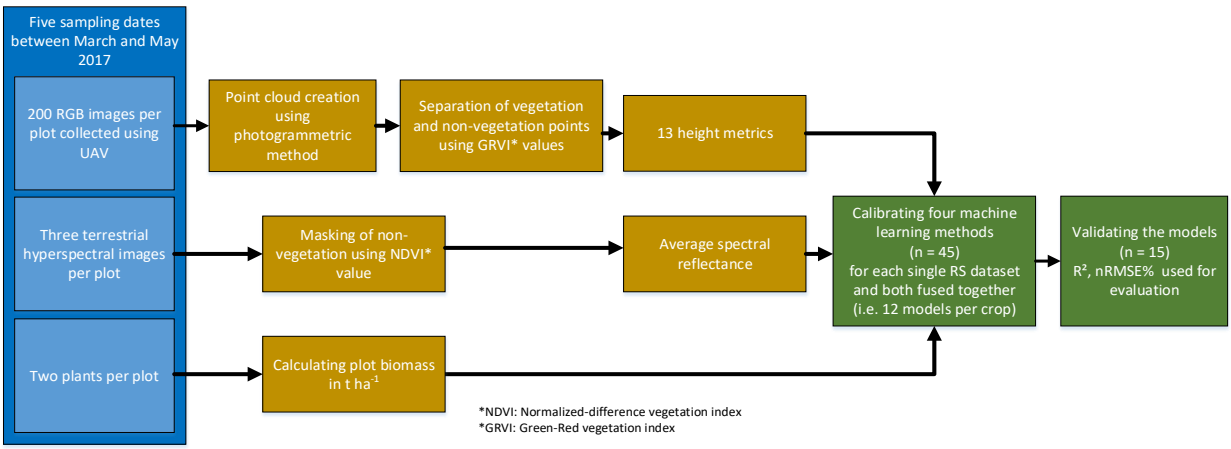


Figure 18: Workflow describing the data collection (blue), data preparation (yellow), and data analysis scheme (green) from multi-factorial irrigated (I) field experiment at University of Agricultural Sciences, GKVK Campus, Bengaluru, India

Sampling of height data

At each sampling date, one RGB UAV flight for each plot was conducted using a DJI FC300X camera (focal length of 3.61 mm) mounted on DJI 3 Professional (DJI, Shenzhen, Guangdong,

China). The same grid-wise flight plan was conducted for each plot using autopilot and the software Pix4D Capture (Pix4D SA, Lausanne, Switzerland). About 200 images were taken during each flight with an 80 % forward and side overlap (Figure 17). For a more detailed description of the flights, see (Moeckel et al., 2018). In total, 60 flights (4 main plots × 3 crops × 5 sampling dates) were conducted. As precise point cloud processing requires high spatial position accuracy, four permanent ground control points (GCPs) were measured in each main plot using a differential GPS (Trimble Inc. Sunnyvale, CA, USA) with a spatial accuracy of ~5 cm. These GCPs were subsequently used for georeferencing, geo-correction, and co-registration of the images. Point cloud data was derived from the UAV RGB images using the structure-from-motion algorithm from the software Photoscan Pro (AgiSoft LLC, St. Petersburg, Russia). The procedures are comprised of a feature extraction process combined with a bundle adjustment of the images (Lowe, 2004; Snavely et al., 2008). Applying a multi-view stereo matching algorithm, a dense point cloud is generated (Mesas-Carrascosa et al., 2015; Westoby et al., 2012). The detailed procedure for the dense point cloud creation is described in Moeckel et al., (2018). For each point of the point cloud, the red and green information was used to calculate the Green-Red Vegetation Index (GRVI; Tucker, 1979). Following the procedure described in Moeckel et al., (2018), the GRVI was used for differentiating points representing ground and vegetation using a k-means classification. With the resulting ground points, a digital elevation model (DEM) was created using a Delaunay triangulation with a spatial resolution of 1 m. The DEM was used to normalise the height information for all points classified as vegetation. The average density of the vegetation point clouds for all plots was 955 points m^{-2}. For each plot and sampling date, 13 point cloud metrics were derived from the vegetation point cloud (Table A supplementary). The selection of the metrics was based on the results from (Silva et al., 2017) and (Næsset and Økland, 2002). In contrast to Moeckel et al., (2018), the minimum crop height parameter was not used for analysis since the minimum height per plot was always zero after conducting height normalisation. The extraction of the point cloud was done using the lidR package (Roussel and Auty, 2017) and the program R (R Core Team, 2016).

Sampling of spectral data

Prior to the biomass sampling, three hyperspectral images were collected in each H-plot using the full-frame hyperspectral camera UHD 185-Firefly (UHD; Cubert GmbH, wwww.cubert-

gmbh.de). Each image was collected above a randomly chosen plant. The distance between the plant and the camera was kept constant at 1.5 m throughout the growing season to keep the area covered by the image during plant growth constant. The camera measures incoming radiance in the spectral range from 450 to 998 nm. However, only the spectral range from 470 to 950 nm was analysed because of too much noise in the wavebands between 450 and 470 nm and 950 and 998 nm. The band width was 4 nm, resulting in 120 bands for which the radiance was measured. The focal length of the camera is 12.1 mm, with an image size of 50 × 50 pixels, resulting in a spatial resolution of 0.02 m and a spectral sampling area of ~1 × 1 m at the applied measurement height. The camera was mounted on a tripod to reduce the effects of shading and movement by the human operator as well as to keep the measurement height constant. Prior to each measurement, the camera was calibrated using a white calibration panel (95% reflectance Zenith Lite), allowing retrieval of reflectance information directly from the measured radiance. During calibration in the field, the integration time (the time over which light is received by the detector) was automatically determined. The reflectance for each image was normalised by dividing the reflectance information for each pixel by the maximum measured reflectance value for this pixel. This normalisation process allowed comparison of spectral information from different sampling dates. Each collected image contained non-vegetation pixels, e.g. soil, shadows, and pipes from the drip irrigation system. To reduce the effect of such disturbances, a two-step procedure was applied (Figure 18). First, the normalised difference vegetation index (NDVI) (Rouse, J. W., Jr. et al., 1974) was calculated for each image. The NDVI is calculated as the difference between reflectance in the red and near-infrared (NIR) spectral range. As red and NIR are characterised by very low (red) and high (NIR) reflectance values for vegetation, they are widely used for differentiating vegetated from non-vegetated surfaces. As second step, a 2-class k-means classification algorithm was applied using the NDVI values. During this process, two class centroids are calculated by measuring the distance from all NDVI values to each other. The centroids are identified in a way that the distance between all pixel values and the corresponding centroids is minimised. Finally, the resulting mask (including pixels of vegetated and non-vegetated surfaces) was used to calculate the average vegetation reflectance for each image and spectral band. All images collected in the same plot and at the same sampling date were averaged. This process results in one spectral reflectance curve per plot and sampling date.

Statistical analysis

A two-factorial ANOVA was used to test for the combined effects of fertilizer and sampling date on the measured biomass. The residuals were checked for normality and homoskedasticity, and base 10 logarithmic transformations were applied as needed to achieve normally distributed residuals.

Supervised ML techniques were used to predict crop biomass for all three crops and identify the best parameter for prediction (height, spectral, or fusion). ML methods are empirical statistical approaches which can be used to establish statistical relationships between a multitude of explanatory variables (here: spectral bands and height parameters) and response variables (here: crop biomass). ML approaches not only allow specification of modelling error and quantification of uncertainty (Holloway and Mengersen, 2018), but also enable the use of highly correlated explanatory variables (Figure 19). Four ML techniques, i.e. partial least squares (PLS), support vector machines (SVM), random forest tree (RFR), and gradient boosting tree (GBT) regression were applied to calibrate the biomass prediction models.

PLS regression searches for a set of latent variables which optimally represent the covariance between the explanatory and response variables. The prediction error of a PLS model decreases with every additional latent variable added to the model until an absolute minimum error is reached. However, this procedure can lead to an overparameterization of the model, resulting in an overoptimistic model. To avoid overparameterization, the search for the optimal number of latent variables was limited to 12, representing 26 % of the total number of samples.

SVM regression is a kernel-based statistical approach. The algorithm uses a kernel function to fit the training data to a hyperplane (Cortes and Vapnik, 1995). The radial basis function was used as a kernel. SVM is particularly suitable for fitting complex non-linear data. SVM models can be optimized by tuning several parameters. In the present study, the optimal model was identified using a systematic search for the best cost and sigma parameters (Karatzoglou et al., 2006).

RFR is a tree-based ensemble learning technique combining the information from many independent decision trees to develop the final prediction model (Breiman, 2001). Determination of the optimal tree required optimization of the parameter mtry, defining how many of the total available explanatory variables will be randomly selected at each split in the regression tree. The number of

trees per iteration was kept constant at 500 in this study. GBT is another ensemble method, but unlike RFR, each tree is trained based on the results of the previous trees (Rich Caruana and Alexandru Niculescu-Mizil, 2006).

For model calibration and validation, the dataset for each crop (n = 60) was split up into a calibration dataset (n = 45) and a validation dataset (n = 15) (Figure 18). The optimisation procedure for the model calibration of all four ML approaches was done using 10-fold group-leave out cross-validation on the calibration dataset. To accomplish this, 40 samples were used for the model calibration, while the remaining five samples were used for model validation. This was repeated ten times, and the optimal parameters for the calibration model were determined subsequently based on the lowest root-mean-square-error (RMSE).

The samples for the validation datasets were selected in a stratified random approach. For each sampling date, three samples of the available 12 samples (3 fertilizer treatments × 4 replicates) were randomly chosen and included in the validation dataset.

To find the best combination of ML approach and variables (i.e. spectral, height, or a fusion of both datasets) for predicting the crop biomass, the explained variance using the R^2_{val} (Equation 1), the prediction error ($RMSE_{val}$) (Equation 2), the relative prediction error (rel.RMSE) (Equation 3), and the bias (Equation 4) were calculated as follows:

$$R^2_{val} = \left[1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y}_i)^2}\right]$$

$$RMSE_{val} = \sqrt{\frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{n}}$$

$$rel.RMSE = \frac{RMSE_{val}}{\max(y_i) - \min(y_i)}$$

$$bias = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)$$

where y_i is the measured biomass, $\hat{y}_i$ is the predicted biomass, $\bar{y}_i$ is the average measured biomass, and n is the number of samples.

Additionally, Taylor diagrams (Taylor, 2001) were used to evaluate the precision of the biomass models. Taylor diagrams provide a quick and easy visual comparison of the predicted versus observed values for several models. For this purpose, the Pearson correlation coefficient (r), the root-mean-square difference (RMSD), and the standard deviation of the predicted and observed values of the validation dataset were calculated for each of the 12 models for each crop.

To examine whether the prediction accuracy of crop biomass is different throughout the growing season, an ANOVA was calculated with the deviation between the predicted and measured biomass values as a response variable and sampling date as an explanatory variable.

Results

The biomass values for the three crops increased throughout the growing season. For eggplant, the range of all biomass values was from 0.013 to 35.62 t ha^{-1}, while biomass yields were lower for tomato (ranging from 0.015 to 19.04 t ha^{-1}) and higher for cabbage (ranging from 0.026 to 65.66 t ha^{-1}). The biomass values for each crop were log10 transformed for the two-factorial ANOVA in order to achieve normally distributed residuals. The results of the ANOVA (Table 6) show that there was no significant effect of fertilizer on biomass, which can be explained by the insufficient range of the fertilizer treatments (Moeckel et al., 2018). The effect of sampling date was significant for all three crops, which was expected as the crops were growing throughout the season. There was no significant interaction effect.

Table 6. Results of the two-factorial ANOVA testing the effect of sampling date (SD) and fertilizer (NF) on the measured biomass of eggplant, tomato, and cabbage from multi-factorial irrigated (I) field experiment at University of Agricultural Sciences, GKVK Campus, Bengaluru, India.

	p-values		
	Sampling Date (SD)	N Fertilizer (NF)	SD × NF
Eggplant	<0.001	0.805	0.670
Tomato	<0.001	0.177	0.575
Cabbage	<0.001	0.594	0.972

Prior to the ML analyses, the intercorrelation between all explanatory variables (i.e. 120 spectral bands and 13 height parameter) was examined. Figure 19 shows a clear intercorrelation between all explanatory variables. However, the pattern of the correlation matrix reveals differences between the three crops. While the wavebands in the near-infrared spectral region (NIR: 770 - 950 nm) are highly correlated with the visible light bands (VIS: 470 -750 nm) for eggplant, this correlation is much weaker for tomato and cabbage. For the 13 height parameters (Table A supplementary), strong correlations (between -0.75 and - 1) were found with bands from the VIS spectral range. The height parameters also showed a strong positive correlation among each other.

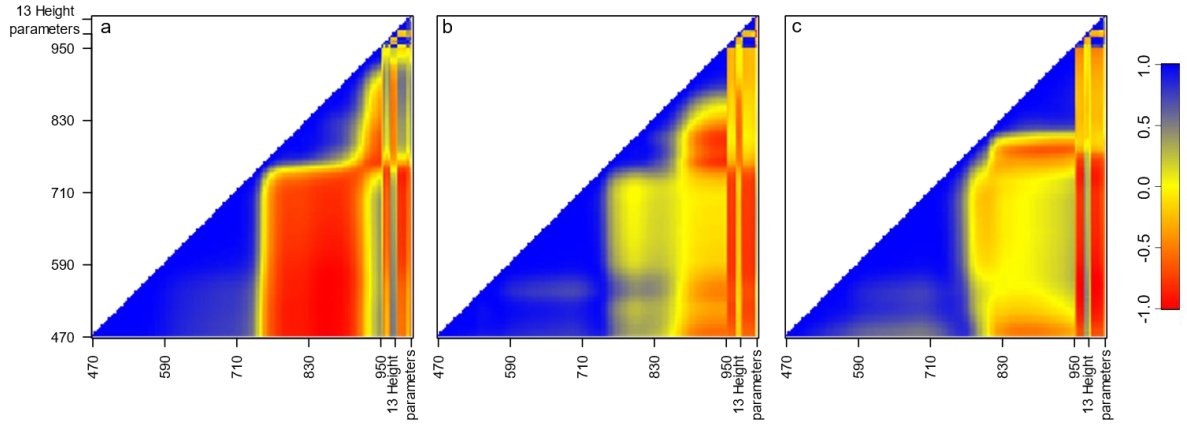


Figure 19: Intercorrelation between all explanatory variables (i.e. 120 spectral bands and 13 height parameters) for a) eggplant, b) tomato, and c) cabbage from multi-factorial irrigated (I) field experiment at University of Agricultural Sciences, GKVK Campus, Bengaluru, India

In total, 36 ML models were calibrated and validated using the test dataset (Figure 20, Table B of the supplementary materials). The prediction accuracy for the three crops varied strongly depending on the calibration model applied and dataset used (eggplant: R^2_{val} = 0.77 - 0.97, nRMSE = 0.07 - 0.18, bias = -0.06 - 0.06; tomato: R^2_{val} = 0.59 - 0.97, nRMSE = 0.06 - 0.23, bias = -0.05 - 0.12; cabbage R^2_{val} = 0.76 - 0.93, nRMSE = 0.09 - 0.17, bias = -0.13 - 0.03) (Figure 5). Only the Taylor diagrams for tomato showed a larger variation of the RMSD (~0.1 - 0.5) and r (~0.8 - 0.99) values for all models (Figure 20B) compared to eggplant and cabbage, indicating dependencies of the selected ML approach and the used dataset for predicting tomato biomass. For eggplant, only PLS regression using the height parameters showed a lower predictive performance (based on RMSD and r) than the other models (Figure 20a). The models predicting cabbage biomass all showed a similar predictive performance (Figure 20 c). For all three crops, RFR showed the best results,

with an R^2_{val} of 0.97 for eggplant and tomato and 0.93 for cabbage (Figure 5). The relative prediction error (nRMSE) was largest for cabbage (9 %), followed by eggplant (7 %) and tomato (6 %). For eggplant, the performance of RFR was only slightly better than the best model of the other methods (R^2_{val} = 0.92, 0.96, and 0.96, and nRMSE = 11 %, 7 %, and 8 % for PLS, SVR, and GBT, respectively). For tomato, the variation in prediction accuracy was the largest, with a R^2_{val} of 0.85, 0.85, and 0.92 and an nRMSE of 14 %, 14 %, and 10 % for the best performing models of PLS, SVR, and GBT, respectively (Table B). For cabbage, the variation in prediction accuracy for the best performing models was low, with an R^2_{val} of 0.91, 0.90, and 0.89 and a relative error of 10 %, 11 %, and 11 % for PLS, SVR, and GBT, respectively.

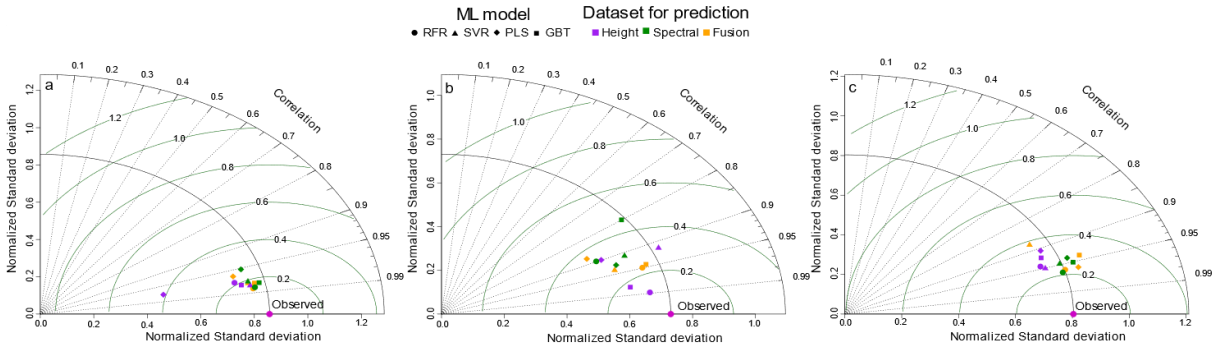


Figure 20. Taylor diagrams displaying a statistical comparison of the 12 biomass estimation models and the observed data of the validation dataset for a) eggplant, b) tomato, and c) cabbage. The best model is located nearest to the observed point (•) on the x-axis while also having a relatively high correlation coefficient (dotted lines) and low RMSD (green lines).

In the following, the evaluation of the three sets of variables (spectral, height, and fusion) is described for the best performing ML models (i.e. RFR) for all crops. The dataset resulting in the most accurate predicted biomass differed between the three crops. For eggplant (Figure 21a, Table B supplementary), the spectral and fusion datasets gave the same prediction accuracy (R^2_{val} = 0.97, nRMSE = 0.07, bias = -0.02), indicating that the height parameters did not improve the prediction model based solely on spectral data. For tomato (Figure 21b), the height dataset (R^2_{val} = 0.97, nRMSE = 0.06, bias = 0.04) clearly outperformed the spectral dataset alone (R^2_{val} = 0.79, nRMSE = 0.16, bias = -0.01) and the fusion of both datasets (R^2_{val} = 0.90, nRMSE = 0.11, bias = 0.01) (Table B). The spectral dataset was best for predicting cabbage biomass (R^2_{val} = 0.93, nRMSE = 0.09, bias = -0.03) (Figure 21c), but did not perform much better than the other two predictor

datasets (R^2_{val} = 0.88, nRMSE = 0.12, bias = -0.10 for the height parameters and R^2_{val} = 0.92, nRMSE = 0.10, bias = -0.04 for the fusion dataset).

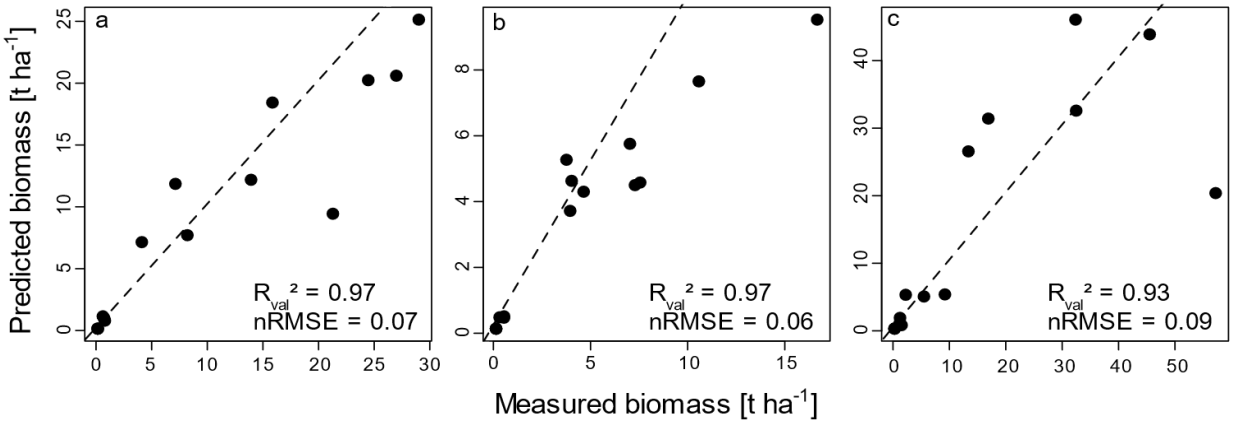


Figure 21. Predicted versus observed biomass based on the best machine learning method (i.e. random forest regression) based on the best performing explanatory set of variables for a) eggplant, predicted using the spectral variables, b) tomato, predicted using the height parameters, and c) cabbage, predicted using the spectral variables from multi-factorial irrigated (I) field experiment at University of Agricultural Sciences, GKVK Campus, Bengaluru, India .

The effect of sampling date on the deviation between predicted and observed biomass (Figure 22), tested by ANOVA, was not significant. However, for eggplant and tomato (Figure 22a and b), an increasing trend towards underestimation of measured biomass is apparent for the later growth stages (i.e. sampling date 4 and 5).

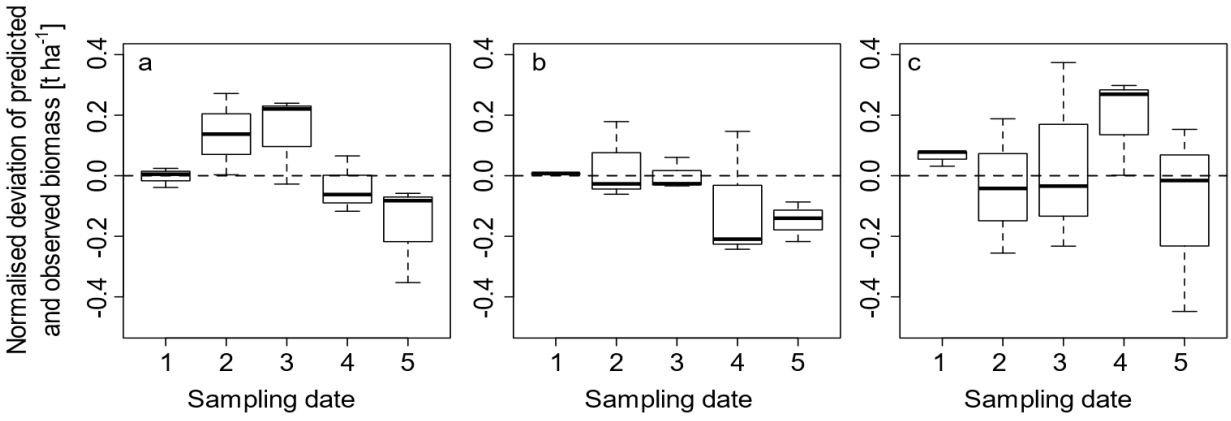


Figure 22. Normalized deviations of the predicted from the observed values for a) eggplant, b) tomato, and c) cabbage for each sampling date. The dotted line indicates zero deviation and thus perfect fit between measured and predicted biomass values.

Discussion

The aim of this study was to develop RS models for predicting fresh biomass of three vegetable crops characterised by differing growth forms and internal vegetation structures. Overall, the results show that high prediction accuracy of fresh biomass can be achieved for vegetable crops with the chosen methods (R^2_{val} of 0.93 for cabbage and 0.97 for tomato and eggplant) (Table B supplementary). The presented direct approach (i.e. predicting biomass directly from remotely sensed variables) resulted in much more accurate biomass estimates in comparison to indirect approaches (Madec et al., 2017; Moeckel et al., 2018). For example, Moeckel et al., (2018) used point cloud-derived height parameters to estimate biomass of the same vegetable crops as in this study via the estimation of the crop height, resulting in a much higher prediction error for all three crops. While Moeckel et al., (2018) obtained the lowest relative prediction error for cabbage compared to the other crops (nRMSE = 15.2 %), the lowest relative prediction error in the present study was obtained for tomato (nRMSE = 6%). Even for cabbage, the relative prediction error in the present study was much lower (nRMSE = 9 %) (Figure 20). Li et al., (2016) tested an approach using point cloud-derived variables and spectral data to predict crop biomass for maize and also achieved lower model accuracies than the present study. The lower accuracies may be partly explained by the fact that their spectral information was based on red, green and blue spectral values without the NIR region, which has been shown to be very important for biomass prediction of vegetation (Bendig et al., 2015).

Regarding the tested ML methods, RFR showed the best prediction performance (based on R^2_{val}, nRMSE, and bias) for all three crops. RFR frequently outperforms other ML methods in predicting biomass from RS data in agriculture (Moeckel et al., 2018; Nasi et al., 2018) and forestry (Powell et al., 2010; Wu et al., 2016). RFR is less sensitive to overfitting than many other methods (Belgiu and Drăguţ, 2016), which might explain the outperformance. Nevertheless, a thorough evaluation of several ML methods should always be conducted, as there is no ML method which consistently outperforms the others under all circumstances (e.g. all ecosystem types and vegetation parameters;Flach, 2012). To evaluate the effect of sampling date on biomass prediction, the deviation of the predicted and measured biomass values for each sampling date was estimated (Figure 22). No significant effect of sampling date on the deviation was found, which does not correspond with results of Malambo et al., (2018) for maize and sorghum and Moeckel et al., (2018) for eggplant, tomato and cabbage, who both found higher deviations for samples from the early crop stages.

Malambo et al., (2018) argued that a discrepancy in sampling time between the biomass and the RS-based sampling date could be the reason for the larger deviations. In Moeckel et al., (2018), the sampling dates for the biomass harvest and the RS sampling were the same and the author argued that the deviation was most likely driven by inaccuracies during the point cloud processing. The slight trend of increasing bias (i.e. biomass underestimation) for the late sampling dates for eggplant and tomato in the present study could be a result of the increased biomass values due to fruit growth, which is no longer related to crop height. Additionally, the reflectance of the entire plant might have changed throughout the growth of the fruits due to increasing water content (Kurunç and Ünlükara, 2009), possibly changing the relationship between biomass and reflectance (Yue et al., 2018). An inclusion of shortwave infrared (>950 nm) spectral information at later growth stages might reduce this bias (Rasooli Sharabian et al., 2013).

For all three crops, the intercorrelation between the explanatory variables (i.e. spectral and height values) are different (Figure 19), particularly in the NIR spectral region, which is known to be strongly related to vegetation biomass and plant internal structure (Knipling, 1970). This could be explained by the different growth forms of the three crops. While eggplant and tomato become tall plants with a dense, closed canopy at the end of growth period, cabbage is characterized by a rather flat structure and open canopy. For tomato, the agricultural practice of tying up the plants with thread might have affected the relationship between biomass and reflectance of NIR. The results of this study therefore confirm the expectation that an individual evaluation of the relationship between biomass and RS variables is necessary for each crop, which has also been shown in e.g. (Malambo et al., 2018; Moeckel et al., 2018; Viña et al., 2011).

To prove whether the fusion of spectral and height information can improve biomass predictions in comparison to single sensor use, a biomass estimation model based on spectral data and point cloud-derived height information alone and in combination was developed. The sensor data fusion did not result in the optimal performance for any of the crops. Although for some of the applied ML approaches, the best model was based on both spectral and height data together (e.g. for the cabbage PLS model, see Table B), the best model for each crop was always based on a single set of variables, i.e. either spectral or height. This could be partly be explained by data redundancy when using both datasets (Pohl and van Genderen, 1998). Wang et al., (2017), who evaluated sensor fusion for crop biomass estimation of maize, found only slight improvements when spectral and lidar data were used together (decrease in prediction error by 0.53 t ha^{-1}), while Prošek and

Šímová, (2019) found a deterioration of prediction accuracy using fused datasets for species identification in shrublands. While for eggplant and cabbage, the spectral dataset delivered the best prediction accuracies, the model based on the height parameters was best for tomato. The tying up of tomato plants from sampling date 3 onwards directly affects the height measurements and thus the point cloud derivation and might be an explanation for the selection of the height parameters. These results confirm the expectation that an evaluation of two or more RS sets of variables reflecting different vegetation characteristics (i.e. spectral reflectance and structure) is as important (Prošek and Šímová, 2019; Tilly et al., 2014) as evaluating different ML methods (Nasi et al., 2018; Powell et al., 2010).

This approach should be further tested using a completely airborne data collection to validate the findings on larger scales and test whether the approach can be used to evaluate spatial variability of crop biomass under field conditions. This step would allow further development of precision agriculture in horticultural systems as it would allow to evaluate phenotypic characteristics of vegetable crops on large scales in a reliable way. The predictive performance of the identified models should be validated with RS data from further growing seasons to check the robustness of the results. Additionally, other sensor systems, such as a thermal imaging system (Maimaitijiang et al., 2017), should be tested for their predictive performance for vegetable crop biomass.

Conclusions

The estimation of vegetable crop fresh matter biomass using in-situ hyperspectral and RGB UAV imagery was successful throughout the growing season. The results also clearly demonstrated that optimal models differ among vegetable crops. Although the fusion of both datasets did not outperform single dataset use, the results indicate that, whenever possible, more than one sensory dataset should be collected to assure an optimal biomass prediction model. Furthermore, it has shown that evaluation of several ML methods using different strategies for building prediction models is worthwhile.

Appendix A

Table A Height parameters derived from point clouds and used as explanatory variables for modelling crop biomass from multi-factorial irrigated (I) field experiment at University of Agricultural Sciences, GKVK Campus, Bengaluru, India. The mean and standard deviation of each metric across all sampling dates and plots is given.

Metric	Description	Mean / Standard deviation
Hmax	Maximum crop height	0.288 / 0.496
Hmean	Mean crop height	0.029 / 0.028
Hsd	Standard deviation of crop height	0.029 / 0.041
Hmedian	Median crop height	0.022 / 0.021
Hskew	Skewness of crop height	2.171 / 2.661
Hkurt	Kurtosis of crop height	19.319 / 97.589
Hcv	Coefficient of variation for crop height	0.978 / 0.843
Hq70	70th percentile of crop height	0.060 / 0.062
Hq80	80th percentile of crop height	0.071 / 0.085
Hq90	90th percentile of crop height	0.088 / 0.110
Hq95	95th percentile of crop height	0.106 / 0.135
Hq99	99th percentile of crop height	0.149 / 0.205

Table B Overview of the prediction performance of the model validation for four ML approaches (PLS: partial least squares, SVM: support vector machines, RFR: random forest regression, and GBT: gradient boosting trees) and three crops (eggplant, tomato, and cabbage). Values in bold represent the model with the best prediction of biomass from multi-factorial irrigated (I) field experiment at University of Agricultural Sciences, GKVK Campus, Bengaluru, India

Crop	Machine learning method used for analysis	Variables used for predicting biomass	R^2	RMSE	nRMSE	Bias
Eggplant	PLS	Height	0.77	0.42	0.18	0.06
		Spectral	0.91	0.26	0.11	0
		Fusion	0.92	0.24	0.11	0.02
	SVM	Height	0.96	0.17	0.07	-0.02
		Spectral	0.95	0.19	0.08	-0.02
		Fusion	0.95	0.18	0.08	-0.02
	RFR	Height	0.94	0.21	0.09	0
		Spectral	**0.97**	**0.15**	**0.07**	**-0.02**
		Fusion	**0.97**	**0.15**	**0.07**	**-0.02**
	GBT	Height	0.95	0.19	0.08	0.03
		Spectral	0.95	0.18	0.08	-0.06
		Fusion	0.96	0.18	0.08	-0.04
Tomato	PLS	Height	0.79	0.34	0.16	0.05
		Spectral	0.85	0.28	0.14	0.02
		Fusion	0.73	0.39	0.19	0.12
	SVM	Height	0.82	0.31	0.15	-0.05
		Spectral	0.82	0.31	0.15	0.06
		Fusion	0.85	0.29	0.14	0.1
	RFR	**Height**	**0.97**	**0.13**	**0.06**	**0.04**
		Spectral	0.79	0.34	0.16	-0.01
		Fusion	0.9	0.23	0.11	0.01
	GBT	Height	0.92	0.2	0.1	0.1
		Spectral	0.59	0.47	0.23	0.12
		Fusion	0.89	0.24	0.12	-0.01
Cabbage	PLS	Height	0.81	0.36	0.15	-0.1
		Spectral	0.87	0.29	0.12	0.03
		Fusion	0.91	0.24	0.1	-0.01
	SVM	Height	0.89	0.26	0.11	-0.09
		Spectral	0.9	0.26	0.11	-0.03
		Fusion	0.76	0.4	0.17	-0.13
	RFR	Height	0.88	0.28	0.12	-0.1
		Spectral	**0.93**	**0.22**	**0.09**	**-0.03**
		Fusion	0.92	0.23	0.1	-0.04
	GBT	Height	0.85	0.32	0.14	-0.08
		Spectral	0.89	0.26	0.11	-0.01
		Fusion	0.86	0.3	0.13	-0.03

References

Aasen, H., & Bolten, A. (2018). Multi-temporal high-resolution imaging spectroscopy with hyperspectral 2D imagers - From theory to application. *Remote Sensing of Environment*, *205*, 374–389.

Atzberger, C. (2013). Advances in remote sensing of agriculture: Context description, existing operational monitoring systems and major information needs. *Remote Sensing*, *5*. https://www.mdpi.com/2072-4292/5/2/949/pdf.

Belgiu, M., & Drăguţ, L. (2016). Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing*, *114*. http://www.sciencedirect.com/science/article/pii/S0924271616000265.

Bendig, J., Yu, K., Aasen, H., Bolten, A., Bennertz, S., Broscheit, J., Gnyp, M.L., & Bareth, G. (2015). Combining UAV-based plant height from crop surface models, visible, and near infrared vegetation indices for biomass monitoring in barley. *International Journal of Applied Earth Observation and Geoinformation*, *39*, 79–87.

Breiman, L. (2001). Random Forests. *Machine Learning*, *45*, 5–32.

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, *20*, 273–297.

Duncan, J.M.A., Dash, J., & Atkinson, P.M. (2015). The potential of satellite-observed crop phenology to enhance yield gap assessments in smallholder landscapes. *Frontiers in Environmental Science*, *3*, 132.

Flach, P. (2012). *Machine learning. The art and science of algorithms that make sense of data.* Cambridge: Cambridge Univ. Press.

Gao, S., Niu, Z., Huang, N., & Hou, X. (2013). Estimating the Leaf Area Index, height and biomass of maize using HJ-1 and RADARSAT-2. *International Journal of Applied Earth Observation and Geoinformation*, *24*. http://www.sciencedirect.com/science/article/pii/S0303243413000172.

Hannerz, F., & Lotsch, A. (2008). Assessment of remotely sensed and statistical inventories of African agricultural fields. *International Journal of Remote Sensing*, *29*, 3787–3804.

Holloway, J., & Mengersen, K. (2018). Statistical machine learning methods and remote sensing for sustainable development goals: A Review. *Remote Sensing*, *10*. https://www.mdpi.com/2072-4292/10/9/1365/pdf.

Karatzoglou, A., Meyer, D., & Hornik, K. (2006). *Support vector machines in R*.

Knipling, E.B. (1970). Physical and physiological basis for the reflectance of visible and near-infrared radiation from vegetation. *Remote Sensing of Environment, 1*. http://www.sciencedirect.com/science/article/pii/S0034425770800219.

Kurunç, A., & Ünlükara, A. (2009). Growth, yield, and water use of okra (Abelmoschus esculentus) and eggplant (Solanum melongena) as influenced by rooting volume. *New Zealand Journal of Crop and Horticultural Science, 37,* 201–210.

Li, W., Niu, Z., Chen, H., Li, D., Wu, M., & Zhao, W. (2016). Remote estimation of canopy height and aboveground biomass of maize using high-resolution stereo images from a low-cost unmanned aerial vehicle system. *Ecological Indicators, 67,* 637–648.

Lowe, D.G. (2004). Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision, 60,* 91–110.

Madec, S., Baret, F., Solan, B. de, Thomas, S., Dutartre, D., Jezequel, S., Hemmerlé, M., Colombeau, G., & Comar, A. (2017). High-throughput phenotyping of plant height: comparing unmanned aerial vehicles and ground LiDAR estimates. *Frontiers in Plant Science, 8.* https://www.frontiersin.org/article/10.3389/fpls.2017.02002.

Maimaitijiang, M., Ghulam, A., Sidike, P., Hartling, S., Maimaitiyiming, M., Peterson, K., Shavers, E., Fishman, J., Peterson, J., Kadam, S., Burken, J., & Fritschi, F. (2017). Unmanned Aerial System (UAS)-based phenotyping of soybean using multi-sensor data fusion and extreme learning machine. *ISPRS Journal of Photogrammetry and Remote Sensing, 134,* 43–58.

Malambo, L., Popescu, S.C., Murray, S.C., Putman, E., Pugh, N.A., Horne, D.W., Richardson, G., Sheridan, R., Rooney, W.L., Avant, R., Vidrinec, M., McCutchen, B., Baltensperger, D., & Bishop, M. (2018). Multitemporal field-based plant height estimation using 3D point

clouds generated from small unmanned aerial systems high-resolution imagery. *International Journal of Applied Earth Observation and Geoinformation*, *64*, 31–42.

Mesas-Carrascosa, F.-J., Torres-Sánchez, J., Clavero-Rumbao, I., García-Ferrer, A., Peña, J.-M., Borra-Serrano, I., & López-Granados, F. (2015). *Assessing optimal flight parameters for generating accurate multispectral orthomosaicks by uav to support site-specific crop management*.

Moeckel, T., Dayananda, S., Nidamanuri, R.R., Nautiyal, S., Hanumaiah, N., Buerkert, A., & Wachendorf, M. (2018). Estimation of vegetable crop parameter by multi-temporal UAV-borne images. *Remote Sensing*, *10*. http://www.mdpi.com/2072-4292/10/5/805.

Mondal, S., Jeganathan, C., Sinha, N.K., Rajan, H., Roy, T., & Kumar, P. (2014). Extracting seasonal cropping patterns using multi-temporal vegetation indices from IRS LISS-III data in Muzaffarpur District of Bihar, India. *The Egyptian Journal of Remote Sensing and Space Science*, *17*. http://www.sciencedirect.com/science/article/pii/S111098231400026X.

Næsset, E., & Økland, T. (2002). Estimating tree height and tree crown properties using airborne scanning laser in a boreal nature reserve. *Remote Sensing of Environment*, *79*, 105–115.

Nasi, R., Viljanen, N., Kaivosoja, J., Alhonoja, K., Hakala, T., Markelin, L., & Honkavaara, E. (2018). Estimating biomass and nitrogen amount of barley and grass using UAV and aircraft based spectral and photogrammetric 3D features. *Remote Sensing*, *10*.

Nidamanuri, R.R., & Zbell, B. (2011). Transferring spectral libraries of canopy reflectance for crop classification using hyperspectral remote sensing data. *Biosystems Engineering*, *110*. http://www.sciencedirect.com/science/article/pii/S1537511011001176.

Padilla, F.M., Gallardo, M., Peña-Fleitas, M.T., Souza, R. de, & Thompson, R.B. (2018). Proximal optical sensors for nitrogen management of vegetable crops: A review. *Sensors*, *18*.

Panigrahy, S., & Sharma, S.A. (1997). Mapping of crop rotation using multidate Indian remote sensing satellite digital data. *ISPRS Journal of Photogrammetry and Remote Sensing*, *52*. http://www.sciencedirect.com/science/article/pii/S0924271697830031.

Pettorelli, N., Vik, J.O., Mysterud, A., Gaillard, J.-M., Tucker, C.J., & Stenseth, N.C. (2005). Using the satellite-derived NDVI to assess ecological responses to environmental change. *Trends in Ecology & Evolution, 20,* 503–510.

Pohl, C., & van Genderen, J.L. (1998). Review article multisensor image fusion in remote sensing: Concepts, methods and applications. *International Journal of Remote Sensing*, *19,* 823–854.

Powell, S.L., Cohen, W.B., Healey, S.P., Kennedy, R.E., Moisen, G.G., Pierce, K.B., & Ohmann, J.L. (2010). Quantification of live aboveground forest biomass dynamics with Landsat time-series and field inventory data: A comparison of empirical modeling approaches. *Remote Sensing of Environment*, *114*. http://www.sciencedirect.com/science/article/pii/S0034425709003745.

Prasad, J.V.N.S., Rao, C.S., Srinivas, K., Jyothi, C.N., Venkateswarlu, B., Ramachandrappa, B.K., Dhanapal, G.N., Ravichandra, K., & Mishra, P.K. (2016). Effect of ten years of reduced tillage and recycling of organic matter on crop yields, soil organic carbon and its fractions in Alfisols of semi arid tropics of southern India. *Soil and Tillage Research*, *156*, 131–139.

Prošek, J., & Šímová, P. (2019). UAV for mapping shrubland vegetation: Does fusion of spectral and vertical information derived from a single sensor increase the classification accuracy? *International Journal of Applied Earth Observation and Geoinformation*, *75*, 151–162.

R Core Team (2016). R. A Language and Environment for Statistical Computing. Vienna, Austria. https://www.R-project.org/.

Rasooli Sharabian, V., Noguchi, N., & Ishi, K. (2013). Optimal vegetation indices for winter wheat growth status based on multi-spectral reflectance. *Environmental Control in Biology*, *51*, 105–112.

Rich Caruana, & Alexandru Niculescu-Mizil (2006). An empirical comparison of supervised learning algorithms using different performance metrics. *Proceedings of the 23rd International Conference on Machine Learning*, 161–168.

Rischbeck, P., Elsayed, S., Mistele, B., Barmeier, G., Heil, K., & Schmidhalter, U. (2016). Data fusion of spectral, thermal and canopy height parameters for improved yield prediction of drought stressed spring barley. *European Journal of Agronomy*, *78*, 44–59.

Rockström, J., Steffen, W., Noone, K., Persson, Å., Chapin III, F.S., Lambin, E.F., Lenton, T.M., Scheffer, M., Folke, C., Schellnhuber, H.J., Nykvist, B., Wit, C.A. de, Hughes, T., van der Leeuw, S., Rodhe, H., Sörlin, S., Snyder, P.K., Costanza, R., Svedin, U., Falkenmark, M., Karlberg, L., Corell, R.W., Fabry, V.J., Hansen, J., Walker, B., Liverman, D., Richardson, K., Crutzen, P., & Foley, J.A. (2009). A safe operating space for humanity. *Nature*, *461*. https://doi.org/10.1038/461472a.

Rouse, J. W., Jr., Haas, R.H., Schell, J.A., & Deering, D.W. (1974). Monitoring vegetation systems in the Great Plains with ERTS. *NASA Special Publication*, *351*, 309.

Roussel, J.-R., & Auty, D. (2017). *lidR: Airborne LiDAR data manipulation and visualization for forestry applications.* https://CRAN.R-project.org/package=lidR.

Silva, C.A., Hudak, A.T., Klauberg, C., Vierling, L.A., Gonzalez-Benecke, C., Padua Chaves Carvalho, S. de, Rodriguez, L.C.E., & Cardil, A. (2017). Combined effect of pulse density and grid cell size on predicting and mapping aboveground carbon in fast-growing Eucalyptus forest plantation using airborne LiDAR data. *Carbon Balance and Management*, *12*, 13.

Snavely, N., Seitz, S.M., & Szeliski, R. (2008). Modeling the world from internet photo collections. *International Journal of Computer Vision*, *80*, 189–210.

Taylor, K.E. (2001). Summarizing multiple aspects of model performance in a single diagram. *Journal of Geophysical Research: Atmospheres*, *106*, 7183–7192.

Thompson, R.B., Voogt, W., Incrocci, L., Fink, M., & Neve, S. de (2018). Strategies for optimal fertiliser management of vegetable crops in Europe. *Acta Horticulturae*, 129–140.

Tilly, N., Hoffmeister, D., Cao, Q., Huang, S., Lenz-Wiedemann, V., Miao, Y., & Bareth, G. (2014). Multitemporal crop surface models. Accurate plant height measurement and biomass estimation with terrestrial laser scanning in paddy rice. *Journal of Applied Remote Sensing*, *8*.

Tilman, D., Balzer, C., Hill, J., & Befort, B.L. (2011). Global food demand and the sustainable intensification of agriculture. *Proceedings of the National Academy of Sciences of the United States of America*, *108*, 20260–20264.

Tucker, C.J. (1979). Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment*, *8*. http://www.sciencedirect.com/science/article/pii/0034425779900130.

UN (2017). World Population Prospects: The 2017 Revision, Key Findings and Advance Tables. Working Paper No. ESA/P/WP/248. United Nations, Department of Economic and Social Affairs, Population Division.

Viña, A., Gitelson, A.A., Nguy-Robertson, A.L., & Peng, Y. (2011). Comparison of different vegetation indices for the remote assessment of green leaf area index of crops. *Remote Sensing of Environment*, *115*. http://www.sciencedirect.com/science/article/pii/S0034425711002926.

Wang, C., Nie, S., Xi, X., Luo, S., & Sun, X. (2017). Estimating the biomass of maize with hyperspectral and LiDAR data. *Remote Sensing*, *9*. https://www.mdpi.com/2072-4292/9/1/11/pdf.

Westoby, M.J., Brasington, J., Glasser, N.F., Hambrey, M.J., & Reynolds, J.M. (2012). 'Structure-from-Motion' photogrammetry: A low-cost, effective tool for geoscience applications. *Geomorphology*, *179*, 300–314.

Wold, S., Sjöström, M., & Eriksson, L. (2008). Partial least squares projections to latent structures (PLS) in chemistry. In P. von Ragu Schleyer, N.L. Allinger, T. Clark, J. Gasteiger, P.A. Kollman, H.F. Schaefer, & P.R. Schreiner (Eds.), *Encyclopedia of Computational Chemistry* (p. 523). [Chichester]: Wiley-Interscience.

Wu, C., Shen, H., Shen, A., Deng, J., Gan, M., Zhu, J., Xu, H., & Wang, K. (2016). Comparison of machine-learning methods for above-ground biomass estimation based on Landsat imagery. *Journal of Applied Remote Sensing*, *10*, 35010.

Yue, J., Feng, H., Yang, G., & Li, Z. (2018). A comparison of regression techniques for estimation of above-ground winter wheat biomass using near-surface spectroscopy. *Remote Sensing*, *10*.

Zhang, C., & Kovacs, J.M. (2012). The application of small unmanned aerial systems for precision agriculture: A review. *Precision Agriculture*, *13*, 693–712.

Zude-Sasse, M., Fountas, S., Gemtos, T.A., & Abu-Khalaf, N. (2016). Applications of precision agriculture in horticultural crops. *European Journal of Horticultural Science*, *81*, 78–90.

Chapter 5

Synthesis

General discussion

The fresh matter biomass (FMB) is one of the vital crop parameters that have been studied by many researchers on different crops (Wang et al. 2017; Yue et al. 2017). In previous studies, different types of RS data have been utilised to estimate crop FMB. However, no study focused on lablab, finger millet and vegetable crop biomass estimation using HS and UAV-borne RGB imagery for an entire cropping season, although they are the major crops grown in the region of Bengaluru. With this background, this thesis aimed to estimate the FMB of some of the cereals, legume and vegetable crops using low altitude remote sensing measurements i.e. UAV-borne RGB, terrestrial HS imaging and sensor data fusion. Biomass was estimated for two cropping seasons, namely monsoon and dry season with different management conditions. Field trials were featured by different levels of N fertiliser and water supply i.e. rainfed and drip-irrigated experiments. Due to multicollinearity of the HS data, ML regression models were employed to handle higher number of predictors. ML methods represent a group of empirical statistical methods which are frequently used for modelling crop biomass from RS data (Maimaitijiang et al., 2017; Moeckel et al., 2018; Rischbeck et al., 2016). Further, various ML methods were utilised for RS based crop biomass estimation and modelling results varied based on the ML method. Therefore, different ML methods PLS, SVM, RFR and GBT were employed for biomass prediction model generation in the current study.

To predict the crop height and assess biomass of vegetable crops using UAV-borne RGB imagery, 14 different crop height metrics were generated from 3D point cloud using the SfM technique for an entire crop growing period (chapter 2). ML methods RFR and SVM were used to predict the crop height. RFR was found to perform best with crop height prediction (cabbage: pseudo-R^2 = 0.97, RMSE = 1.3, bias = 0.01, eggplant: pseudo-R^2 = 0.93, RMSE = 6.86, bias = 0.14, tomato: pseudo-R^2 = 0.89, RMSE = 5.49, bias = 0.45). It was found that crop height for vegetable crops could be estimated with R^2 ranging from 0.89 for tomato to 0.97 for cabbage with different crop growth forms. These results are similar to multi-temporal crop height models for maize and sorghum (R^2 ranging from 0.68 to 0.78) (Malambo et al. 2018).

Along the crop growth stages there was a systematic deviation of the median from observed and predicted crop height. There was an increase in the median deviation values from early to late growth stage of the crop in cabbage and eggplant. Another study (Malambo et al. 2018) with maize and sorghum reported similar findings. In the cited study, there was a time lag between the UAV measurements and the field measurements, while in our study, there was no time difference between these measurements. Hence, the biased estimation of height might be from inaccuracies in the point cloud generation process. For tomato, there was high bias values at sampling date 4, as the plants were tied horizontally with the threads from sampling date 3. Plants were tied as an agricultural practice to prevent the fruits from moulding.

The predicted crop height from RFR was used as an estimator of biomass using linear regression in chapter 2. Cabbage, tomato and eggplant are featured by distinctively different growth structures. Cabbage is round and ball structured while being attached to the ground, tomato is more tender and delicate in growth and hence tied to strings for support. Eggplant is featured by dense growth with broad leaves and thick stem. Crop height is considered to be an essential indicator of biomass estimation in barley (Bendig et al. 2014), sorghum (Malambo et al. 2018) and poppy (Iqbal et al. 2017). The predicted crop height shows a strong and significant relationship with the biomass of all the three crops investigated in this study (chapter 2). The prediction accuracy of biomass was better in the present study than in previous studies (Garcia-Gutierrez et al. 2014), where UAV-borne RGB imagery was used to estimate maize biomass. Additionally, the current study comprised multi temporal biomass measurements at 20 m flying height, compared to 130 m flying height at a single date measurement (Li et al. 2016). Prediction accuracies of crop height derived biomass showed similar prediction qualities between predicted and field measured crop heights. This leads to the conclusion that RS can estimate the biomass of the vegetable crops eggplant, tomato and cabbage using UAV-borne RGB images. The analysis approach based on the extraction of several height parameter from 3D point clouds and ML regression methods showed a stable model prediction quality for biomass during the entire growing period. Therefore, time consuming intensive crop height measurements can be replaced by point cloud analysis which can provide a similar prediction accuracy for biomass.

The modelling approach for height estimation of vegetable crops showed an increasing deviation during late growth stages (growth stage 4 and 5). This deviation indicates that crop height is a less

reliable indicator of biomass during late growth stages. The height of the crops stops to grow vertically at a certain stage and starts growing horizontally with higher stem thickness and broadening of leaves. Hence, fusing various RS data might compensate for the weakness of an individual sensor data.

Monsoon crop biomass (lablab, maize and finger millet) was estimated using terrestrial HS images. Treatments covered different N levels and water supply. Analysis was conducted using ML method RFR. The study covered a wide range of phenological stages during three consecutive cropping seasons in year 2016 (Y1), 2017 (Y2) and 2018 (Y3). Hence, it is an important aspect considering the validity range of generalised models with different levels of N fertiliser and water supply. The generalised models had the lowest prediction error (with rRMSEP ranging from 13.9 % to 18.7 %) when compared to rainfed and irrigated experiments data models seperately. The range of crop productivity became much higher with a combination of data from two different water conditions, rainfed and irrigated experiments. Integration might have increased the robustness of the regression models. Similar prediction errors (relative error 16.7%, $R^2 = 0.78$) were found in (Li et al. 2016) for maize biomass modelled from RGB images. In contrast to our study, Li et al. (2016) included canopy height parameter additional to RGB information. It shows the potential of structural data calculated with photogrammetric methods (Bendig et al. 2014; Moeckel et al. 2018). The highest error was found for the productive crops maize and finger millet. Similarly, finger millet is also a rarely researched crop in the field of RS. The second hypothesis, that terrestrial HS imaging can be used to estimate the biomass of monsoon crops lablab, maize and finger millet, could be verified. Generalised models which were built on both rainfed and irrigated conditions, were more robust than the individual water supply specific models.

The normalised deviation of median values from predicted and observed FMB values, based on N levels and water supply was close to zero. It shows the model is more robust, as the prediction of FMB is independent of N levels and water supply. Based on sampling dates, there was a particular pattern observed in the estimation of biomass in Y1 and Y2 points. Although the exclusion of soil spectral signatures was done, there was overestimation of biomass at the initial sampling dates of lablab, maize and finger millet. The overestimation might have resulted from the presence of weeds at the initial sampling dates. At the later stages, it was found there was an underestimation of biomass and the relative error increased in the order of lablab, finger millet and maize with the

increasing height of the crop. This indicates that higher parts of the plants masked reflectance information from lower parts of the plants. This effect is known as "saturation constraint" and was regularly found in previous studies (Baez-Gonzalez et al. 2002; Freeman et al. 2007; Pimstein, Karnieli, and Bonfil 2007) mainly when NDVI was used. The increase in the relative error with increasing height of the plants and underestimation of biomass at later stages stresses the multi sensor approach which may overcome the shortcomings of individual sensors. According to our hypothesis, this proves that the normalised deviation of predicted and observed biomass, varies based on sampling dates in monsoon crops. But with different N levels and water supply, normalised deviation did not vary much contradicting the hypothesis. It could be shown that the generalised model can estimate the biomass irrespective of different N levels and water supply.

The important spectral bands affecting the estimation of biomass were found in the red-edge region of the electromagnetic spectrum for lablab, maize and finger millet. In the generalised model, lablab had several important wavelengths in the green, red and NIR region indicating a large number of important variables in these models. Similar vital bands were found by (Manjunath, Ray, and Panigrahy 2011) in the discrimination of pulses like chickpea, pea and lentils. However, in maize and finger millet the most important variables were found in the red-edge region and the red region. The ability of the leguminous lablab to fix atmospheric nitrogen may have resulted in a longer greenness of the crop during the growing period. This leads to higher reflectance at the green peak (~550nm) and higher absorbance in red (~650 nm). In general, the identified spectral bands confirm the accepted knowledge about the biomass reflectance relationships.

As an individual sensor was used for biomass estimation of the vegetable crops in chapter 2, sensor data fusion of UAV-borne RGB and terrestrial HS measurements were studied in chapter 4 to further improve the prediction accuracy. The objective was to test the predictive performance of these two RS data in FMB estimation for three vegetable crops (eggplant, tomato and cabbage). Out of the four different ML methods, RFR showed the best prediction accuracy (based on R^2_{val} and nRMSE) for all three crops. In the present study the lowest relative error was found for tomato (nRMSE = 6%) from height variable used for predicting biomass. However, for the same vegetable crops in chapter 2, the lowest relative error was found for cabbage (nRMSE = 15.2%) when compared to other crops. Even for cabbage in chapter 4, the relative prediction error was much lower (nRMSE = 9%) from spectral variables. For the estimation of biomass, the sensor data fusion did

not outperform the individual sensors among all three vegetable crops. Additionally, it seems to be better to predict the FMB directly from the remotely sensed variables instead of using indirect approaches (Madec et al. 2017; Moeckel et al. 2018). As point cloud derived crop height modelling to estimate the biomass in chapter 2 resulted in a much higher prediction error for all three crops. Confounding our third hypothesis, sensor data fusion did not improve the prediction accuracies when compared to the use of an individual sensor, but each single sensor performed best in different vegetable crops. Hence, different crop parameter measuring sensors might help to improve prediction accuracies of biomass.

The inter-correlation between the spectral and height values were different particularly in the NIR region which is noticed to be strongly dependent on biomass and internal structures of the crop. As all these three vegetable crops have different growth forms, eggplant and tomato grow with height and form a dense canopy while cabbage is more flat and open. In tomato, the management practice of tying the plants with the thread might have affected the relationship of biomass and NIR reflectance.

Based on the tested ML methods, RFR showed the best prediction performance for all three vegetable crops in chapter 4. The relative prediction error (nRMSE) was largest for cabbage (9 %), followed by eggplant (7 %) and tomato (6 %). For eggplant, the performance of RFR was only slightly better than the best model of the other methods (nRMSE = 11 %, 7 %, and 8 % for PLS, SVR, and GBT, respectively). For tomato, the variation in prediction error was the largest, with a nRMSE of 14 %, 14 % and 10 % for the best performing models of PLS, SVR, and GBT, respectively. For cabbage, the variation in prediction error for the best performing models was low, with a relative error of 10 %, 11 %, and 11 % for PLS, SVR, and GBT, respectively. RFR frequently outperformed the other ML method in agricultural applications (Moeckel et al. 2018; Nasi et al. 2018) and is less sensitive to overfitting than many other methods which might explain the best performance (Belgiu and Drăguţ 2016). However, a thorough evaluation of ML methods should be conducted, as there is no ML method which consistently outperforms the other method under all circumstances (e.g. all ecosystem and vegetation parameters;(Flach 2012).

There was an underestimation of biomass at the later growth stages (sampling date 4 and 5) from the best models of RFR for eggplant, tomato and cabbage (chapter 4). This might have resulted from the increase in the biomass values and fruits size which is no longer related to the crop height.

Also, the reflectance of the entire plant might have changed due to the fruit growth and increase in the water content. It might alter the relationship between biomass and spectral reflectance (Yue, Feng, Yang, et al. 2018). Inclusion of shortwave infrared (>950 nm) spectral information at the later growth stages might reduce the bias (Rasooli Sharabian, Noguchi, and Ishi 2013). According to our hypothesis, the variation based on sampling dates was found for vegetable crops. However, for eggplant and tomato, an increasing trend of underestimation of normalised deviation of predicted and observed biomass was found for later crop growth stages.

The developed FMB models for different cereals, legume and vegetable crops from the experimental fields covered a wide range of crop phenology. As these crops are widely grown around the region of Bengaluru, the FMB models developed in experimental plots have to be validated in larger areas of field conditions to determine the robustness of the model. Terrestrial HS measurements are time consuming and cannot be applied to large areas. Therefore, upscaling HS sensor with UAV might be a solution to measure the reflectance of crop canopies in larger areas.

In the study of (Rama Rao, Garg, and Ghosh 2007), the possible integration of in-situ HS measurements with space borne HS RS data for automatic identification and discrimination of various crop cultivars using crop specific spectral libraries was examined. Similarly, the terrestrial HS measurements can be used as reference data for identification and classification of crops from satellite images in larger area. Hereby, possible land use changes and socio-economic effects on the cropping patterns can be analysed. in rural urban gradient of Bengaluru.

The FMB models developed in this thesis can be used by the implementing agencies of the government for yield estimation at field level. As the complete growing period of crops were covered, the current study might improve estimation of the yield already during early stages and enable interventions for yield improvement. The advanced RS technology with market forces may help in speculating the demand-supply and in stabilising the price fluctuations. It thereby might safeguard the farmers from a price crash and provide secure income. Additionally, crop insurance companies can make use of the FMB models through providing compensations based on the estimated yield (Shirsath et al. 2019). When farming becomes profitable and the risk involved is reduced, it may reduce the rural youth from migrating to cities in search of employment and thereby reducing the rural-urban transformation. This study is applicable to one of the biggest transforming cities - Bengaluru - where rural-urban transformation is taking place at a faster pace. The outcome

of the studies in the thesis can be integrated with other disciplines of ecology, economics and social sciences for a holistic approach towards analysing patterns of change based on variability at different spatial scales in the rural-urban gradient.

Chapter 6

Summary

India, the agriculture based country with multiple cropping seasons and crop diversity is witnessing a rapid rural-urban transformation. To feed the growing population, various agro-technological advancements with crop intensification are of much interest. Therefore, with the aim of transferring technology from experimental fields to the practical farmers' fields, estimation of biomass of major crops grown in Bengaluru were studied with terrestrial HS and UAV-borne RGB RS under various management levels.

(i) The estimation of vegetable crop biomass using UAV-borne RGB 3D point clouds showed similar model prediction qualities in biomass estimation between measured and predicted crop heights. Hence, the method offers an alternative to time consuming height measurements and the need of real-time evaluation of vegetable biomass.

(ii) The multi-temporal biomass estimation of cereals and legume crops using terrestrial HS images had less relative errors for lablab followed by finger millet and maize. This indicates the requirement of an additional sensor which can overcome declining model prediction accuracy with increased plant height.

(iii) The sensor data fusion did not improve the prediction accuracies when compared to data from an individual sensor. However, every single sensor performed best for different vegetable crops. Hence, multi sensor data has to be collected to improve the biomass estimation of different crops.

(iv) The estimation of biomass did not vary much for cereals and legume monsoon crops based on N levels and water supply. Whereas, there was overestimation in the initial sampling dates and underestimation at final sampling dates for lablab, finger millet and maize. From the best RFR model of two individual sensors in vegetable crops, an increasing underestimation was found for eggplant, tomato and cabbage.

(v) Out of the various ML methods, RFR was found to be outperforming in height and biomass estimation of all the crops in the study.

(vi) The important spectral bands in the estimation of cereals and legume monsoon biomass in generalised models were found in the red-edge region for all the three crops. However, several other wavelengths in the visible and near-infrared region were important in models for lablab and finger millet.

With the evaluation of different remote sensing based spectral and 3D point cloud data for crop biomass estimation in southern India, the study can be further upscaled to UAV and spaceborne images to cover large areas of changing cropping pattern in the rural-urban gradient.

Chapter 7

Zusammenfassung

Indien, ein Land geprägt durch eine Landwirtschaft mit mehrfachen Hauptkulturen pro Jahr und einer hohen Kulturpflanzen-Diversität, erlebt eine rapide ländlich-urbane Transformation. Mit dem Bedarf der Nahrungsmittelversorgung einer wachsenden Bevölkerung steigt das Interesse an agro-technologischen Fortschritten und der Intensivierung der Anbauverfahren. Mit dem Ziel, technologische Erkenntnisse aus Feldexperimenten an Landwirte zu transferieren, wurde der Ertrag verschiedener in Bengaluru angebauter Hauptkulturen mittels eines terrestrisch, hyperspektralen- und eines drohnenbasierten RGB-Fernerkundungssystems unter verschiedenen Bewirtschaftungsformen untersucht.

(i) Die Ertragsabschätzung aus 3D-Punktwolken, erzeugt aus drohnenbasierten RGB-Bildern, erzielte vergleichbare Schätzgenauigkeiten, wie Biomasseschätzungen mittels im Feld erhobener Vegetationshöhen. So stellt diese Methode eine Alternative zu zeitaufwändigen Höhenmessungen im Feld dar und liefert eine Echtzeit-Evaluierung von Feldfrucht-Erträgen.

(ii) Die multitemporale Ertragsabschätzung von Getreide und Gemüse mittels terrestrischer Hyperspektralbilder erzielte den geringsten relativen Fehler für LabLab, gefolgt von Fingerhirse und Mais. Dies zeigt, dass für Ertragsabschätzungen von hochwachsenden Kulturen der Bedarf nach alternativen Sensorsystem steigt.

(iii) Die Sensor-Fusionierung erbrachte keine Verbesserung in der Ertragsabschätzung, verglichen mit Modellen aus einzelnen Sensordaten. Diese Einzelsensor-Modelle zeigten ihre Stärken vor allem in der Abschätzung einzelner Kulturarten. Für eine Abschätzung mehrerer Kulturarten gilt es jedoch, die Sensordaten zu kombinieren.

(iv) Bei den Getreide- und Monsun-Gemüse-Kulturen variierte die Biomasseabschätzung nicht sonderlich hinsichtlich der N-Düngergabe und Bewässerungsvariante. Die Proben des ersten Messtermins wurden jedoch überschätzt und die des letzten Messtermins unterschätzt für LabLab, Fingerhirse und Mais. Das beste RFR-Modell aus Einzelsensor-

Daten zeigte eine zunehmende Unterschätzung der Biomassewerte für Aubergine, Tomate und Kohl.

(v) Aus zahlreichen Methoden des maschinellen Lernens erzielte die RFR die höchsten Schätzgenauigkeiten zur Höhen- und Biomasseabschätzung für alle Kulturen.

(vi) Die wichtigsten Wellenlängenbereiche zur Schätzung von Getreide- und Gemüseerträgen befanden sich im Red-Edge-Bereich für alle drei Kulturen. Jedoch waren auch andere Wellenlängen aus dem sichtbaren und nah-infraroten Bereich wichtig bei der Modellierung der Biomasse von Lablab und Finger Hirse.

Die Ansätze aus spektralen und 3D-Punktwolken-Daten zur Biomasseabschätzung in Süd-Indien können durch drohnen- und sattelitenbasierte Plattformen erweitert und für Aufnahmen von sich veränderndem Kulturpflanzenanbau in ländlich-urbanen Gradienten eingesetzt werden.

Chapter 8

References

Aasen, Helge, and Andreas Bolten. 2018. "Multi-Temporal High-Resolution Imaging Spectroscopy with Hyperspectral 2D Imagers - From Theory to Application." *Remote Sensing of Environment* 205: 374–89.

Aasen, Helge, Andreas Burkart, Andreas Bolten, and Georg Bareth. 2015. "Generating 3D Hyperspectral Information with Lightweight UAV Snapshot Cameras for Vegetation Monitoring: From Camera Calibration to Quality Assurance." *ISPRS Journal of Photogrammetry and Remote Sensing* 108: 245–59. https://www.sciencedirect.com/science/article/pii/S0924271615001938?via%3Dihub (May 6, 2019).

Aasen, Helge, Martin Leon Gnyp, Yuxin Miao, and Georg Bareth. 2014. "Automated Hyperspectral Vegetation Index Retrieval from Multiple Correlation Matrices with HyperCor." *Photogrammetric Engineering & Remote Sensing* 80(8): 785–95.

Arjun, Kekane Maruti. 2013. "Indian Agriculture-Status, Importance and Role in Indian Economy." *International Journal of Agriculture and Food Science Technology* 4(4): 343–46. http://www.ripublication.com/ijafst.htm (May 14, 2019).

Baez-Gonzalez, Alma Delia, Pei-yu Chen, Mario Tiscareno-Lopez, and Srinivasan Raghavan. 2002. "Using Satellite and Field Data with Crop Growth Modeling to Monitor and Estimate Corn Yield in Mexico." *Crop Science* 42: 1943–49.

Belgiu, Mariana, and Lucian Drăguţ. 2016. "Random Forest in Remote Sensing: A Review of Applications and Future Directions." *ISPRS Journal of Photogrammetry and Remote Sensing* 114: 24–31. http://www.sciencedirect.com/science/article/pii/S0924271616000265.

Bendig, Juliane et al. 2014. "Estimating Biomass of Barley Using Crop Surface Models (CSMs) Derived from UAV-Based RGB Imaging." *Remote Sensing* 6(11): 10395–412. http://www.mdpi.com/2072-4292/6/11/10395 (June 17, 2019).

Burkart, Andreas et al. 2015. "Angular Dependency of Hyperspectral Measurements over Wheat Characterized by a Novel UAV Based Goniometer." *Remote Sensing* 7(1): 725–46. http://www.mdpi.com/2072-4292/7/1/725 (May 6, 2019).

Dayananda, Supriya et al. 2019. "Multi-Temporal Monsoon Crop Biomass Estimation Using Hyperspectral Imaging." *Remote Sensing* 11(15): 1771. https://www.mdpi.com/2072-4292/11/15/1771 (July 30, 2019).

Ferrant, Sylvain et al. 2017. "Detection of Irrigated Crops from Sentinel-1 and Sentinel-2 Data to Estimate Seasonal Groundwater Use in South India." *Remote Sensing* 9(11): 1119. http://www.mdpi.com/2072-4292/9/11/1119 (April 29, 2019).

Flach, Peter. 2012. *Machine Learning: The Art and Science of Algorithms That Make Sense of Data*. Cambridge: Cambridge Univ. Press.

Freeman, Kyle W. et al. 2007. "By-Plant Prediction of Corn Forage Biomass and Nitrogen Uptake at Various Growth Stages Using Remote Sensing and Plant Height." *Agronomy Journal* 99(2): 530–36.

Gao, Shuai, Zheng Niu, Ni Huang, and Xuehui Hou. 2013. "Estimating the Leaf Area Index, Height and Biomass of Maize Using HJ-1 and RADARSAT-2." *International Journal of Applied Earth Observation and Geoinformation* 24: 1–8. http://www.sciencedirect.com/science/article/pii/S0303243413000172.

Garcia-Gutierrez, Jorge, Francisco Martínez-Álvarez, Alicia Troncoso, and Jose C Riquelme. 2014. "A Comparative Study of Machine Learning Regression Methods on LiDAR Data: A Case Study." In *International Joint Conference SOCO'13-CISIS'13-ICEUTE'13*, eds. Álvaro Herrero et al. Cham: Springer International Publishing, 249–58.

Grüner, Esther, Thomas Astor, and Michael Wachendorf. 2019. "Biomass Prediction of Heterogeneous Temperate Grasslands Using an SfM Approach Based on UAV Imaging." *Agronomy* 9(2): 54. http://www.mdpi.com/2073-4395/9/2/54 (September 11, 2019).

Hannerz, F, and A Lotsch. 2008. "Assessment of Remotely Sensed and Statistical Inventories of African Agricultural Fields." *International Journal of Remote Sensing* 29(13): 3787–3804.

Honkavaara, Eija et al. 2013. "Processing and Assessment of Spectrometric, Stereoscopic Imagery Collected Using a Lightweight UAV Spectral Camera for Precision Agriculture." *Remote Sensing* 5(10): 5006–39. http://www.mdpi.com/2072-4292/5/10/5006 (May 6, 2019)."India's 10 Fastest Growing Cities." 2008. *Rediff News*.

http://us.rediff.com/money/2008/feb/06sld2.htm (May 7, 2019).

Iqbal, Faheem, Arko Lucieer, Karen Barry, and Reuben Wells. 2017. "Poppy Crop Height and Capsule Volume Estimation from a Single UAS Flight." *Remote Sensing* 9(7).

Krishna, Gopal et al. 2019. "Comparison of Various Modelling Approaches for Water Deficit Stress Monitoring in Rice Crop through Hyperspectral Remote Sensing." *Agricultural Water Management* 213(August 2018): 231–44. https://doi.org/10.1016/j.agwat.2018.08.029.

Lati, Ran Nisim, Sagi Filin, and Hanan Eizenberg. 2013. "Estimating Plant Growth Parameters Using an Energy Minimization-Based Stereovision Model." *Computers and Electronics in Agriculture* 98: 260–71. http://www.sciencedirect.com/science/article/pii/S0168169913001622.

Li, Wang et al. 2016. "Remote Estimation of Canopy Height and Aboveground Biomass of Maize Using High-Resolution Stereo Images from a Low-Cost Unmanned Aerial Vehicle System." *Ecological Indicators* 67: 637–48.
https://www.sciencedirect.com/science/article/pii/S1470160X16301406?via%3Dihub (May 31, 2019).

Li, Wang, Zheng Niu, Hanyue Chen, and Dong Li. 2017. "Characterizing Canopy Structural Complexity for the Estimation of Maize LAI Based on ALS Data and UAV Stereo Images." *International Journal of Remote Sensing* 38(8–10): 2106–16.

Madec, Simon et al. 2017. "High-Throughput Phenotyping of Plant Height: Comparing Unmanned Aerial Vehicles and Ground LiDAR Estimates." *Frontiers in Plant Science* 8: 2002. https://www.frontiersin.org/article/10.3389/fpls.2017.02002.

Maimaitijiang, Maitiniyazi et al. 2017. "Unmanned Aerial System (UAS)-Based Phenotyping of Soybean Using Multi-Sensor Data Fusion and Extreme Learning Machine." *ISPRS Journal of Photogrammetry and Remote Sensing* 134: 43–58.
https://linkinghub.elsevier.com/retrieve/pii/S0924271617303246.

Malambo, L et al. 2018. "Multitemporal Field-Based Plant Height Estimation Using 3D Point Clouds Generated from Small Unmanned Aerial Systems High-Resolution Imagery." *International Journal of Applied Earth Observation and Geoinformation* 64: 31–42. https://linkinghub.elsevier.com/retrieve/pii/S0303243417301800.

Manjunath, K. R., Shibendu Shankar Ray, and Sushma Panigrahy. 2011. "Discrimination of

Spectrally-Close Crops Using Ground-Based Hyperspectral Data." *Journal of the Indian Society of Remote Sensing* 39(4): 599–602. http://link.springer.com/10.1007/s12524-011-0099-x (June 5, 2019).

Moeckel, Thomas et al. 2018. "Estimation of Vegetable Crop Parameter by Multi-Temporal UAV-Borne Images." *Remote Sensing* 10(5): 805. http://www.mdpi.com/2072-4292/10/5/805 (April 3, 2019).

Mondal, Saptarshi et al. 2014. "Extracting Seasonal Cropping Patterns Using Multi-Temporal Vegetation Indices from IRS LISS-III Data in Muzaffarpur District of Bihar, India." *The Egyptian Journal of Remote Sensing and Space Science* 17(2): 123–34. http://www.sciencedirect.com/science/article/pii/S111098231400026X.

Nasi, Roope et al. 2018. "Estimating Biomass and Nitrogen Amount of Barley and Grass Using UAV and Aircraft Based Spectral and Photogrammetric 3D Features." *Remote Sensing* 10(7).

Nidamanuri, Rama Rao, and Bernd Zbell. 2011. "Transferring Spectral Libraries of Canopy Reflectance for Crop Classification Using Hyperspectral Remote Sensing Data." *Biosystems Engineering* 110(3): 231–46. http://www.sciencedirect.com/science/article/pii/S1537511011001176.

Panigrahy, S, and S A Sharma. 1997. "Mapping of Crop Rotation Using Multidate Indian Remote Sensing Satellite Digital Data." *ISPRS Journal of Photogrammetry and Remote Sensing* 52(2): 85–91. http://www.sciencedirect.com/science/article/pii/S0924271697830031.

Park, Suyoung et al. 2017. "Adaptive Estimation of Crop Water Stress in Nectarine and Peach Orchards Using High-Resolution Imagery from an Unmanned Aerial Vehicle (UAV)." *Remote Sensing* 9(8): 828. http://www.mdpi.com/2072-4292/9/8/828.

Parkes, Stephen D, Matthew F McCabe, Samir K Al-Mashhawari, and Jorge Rosas. 2016. "Reproducibility of Crop Surface Maps Extracted from Unmanned Aerial Vehicle (UAV) Derived Digital Surface Maps." In *Remote Sensing for Agriculture, Ecosystems, and Hydrology XVIII*, Proceedings of SPIE, eds. C M U Neale and A Maltese.

Pimstein, Agustin, Arnon Karnieli, and David J Bonfil. 2007. "Wheat and Maize Monitoring

Based on Ground Spectral Measurements and Multivariate Data Analysis." *Journal of Applied Remote Sensing* 1(1): 013530.

Pohl, C, and J L van Genderen. 1998. "Review Article Multisensor Image Fusion in Remote Sensing: Concepts, Methods and Applications." *International Journal of Remote Sensing* 19(5): 823–54.

Rama Rao, N, P.K. Garg, and S.K. Ghosh. 2007. "Development of an Agricultural Crops Spectral Library and Classification of Crops at Cultivar Level Using Hyperspectral Data." *Precision Agriculture* 8(4–5): 173–85. http://link.springer.com/10.1007/s11119-007-9037-x (June 27, 2019).

Rasooli Sharabian, Vali, Noboru Noguchi, and Kazunobu Ishi. 2013. "Optimal Vegetation Indices for Winter Wheat Growth Status Based on Multi-Spectral Reflectance." *Environmental Control in Biology* 51(3): 105–12.

Rischbeck, Pablo et al. 2016. "Data Fusion of Spectral, Thermal and Canopy Height Parameters for Improved Yield Prediction of Drought Stressed Spring Barley." *European Journal of Agronomy* 78: 44–59. https://linkinghub.elsevier.com/retrieve/pii/S1161030116300752.

Rouse, J W, R H Haas, Schell J.A., and D W Deering. 1974. "Monitoring Vegetation Systems in the Great Plains with ERTS." *NASA Special Publication* 1(Sect. A): 309–17. https://ntrs.nasa.gov/search.jsp?R=19740022614 (March 27, 2019).

Schirrmann, Michael et al. 2016. "Monitoring Agronomic Parameters of Winter Wheat Crops with Low-Cost UAV Imagery." *Remote Sensing* 8(9).

Selsam, Peter, Wolfgang Schaeper, Katja Brinkmann, and Andreas Buerkert. 2017. "Acquisition and Automated Rectification of High-Resolution RGB and near-IR Aerial Photographs to Estimate Plant Biomass and Surface Topography in Arid Agro-Ecosystems." *Experimental Agriculture* 53(1): 144–57.

Shirsath, Paresh, Shalika Vyas, Pramod Aggarwal, and Kolli N. Rao. 2019. "Designing Weather Index Insurance of Crops for the Increased Satisfaction of Farmers, Industry and the Government." *Climate Risk Management* 25: 100189. https://www.sciencedirect.com/science/article/pii/S2212096318300998 (August 29, 2019).

Thenkabail, Prasad et al. 2009. "Irrigated Area Maps and Statistics of India Using Remote Sensing and National Statistics." *Remote Sensing* 1(2): 50–67. http://www.mdpi.com/2072-4292/1/2/50 (May 2, 2019).

United Nations. 2018. *World Urbanization Prospects: The 2018 Revision*. https://population.un.org/wup/Publications/Files/WUP2018-KeyFacts.pdf (April 25, 2019).

Vigneau, Nathalie, Martin Ecarnot, Gilles Rabatel, and Pierre Roumet. 2011. "Potential of Field Hyperspectral Imaging as a Non Destructive Method to Assess Leaf Nitrogen Content in Wheat." *Field Crops Research* 122(1): 25–31. https://www.sciencedirect.com/science/article/pii/S0378429011000451 (May 6, 2019).

Wang, Cheng et al. 2017. "Estimating the Biomass of Maize with Hyperspectral and LiDAR Data." *Remote Sensing* 9(1): 1–12.

Warren, Georgina, and Graciela Metternicht. 2005. "Agricultural Applications of High-Resolution Digital Multispectral Imagery: Evaluating within-Field Spatial Variability of Canola (Brassica Napus) in Western Australia." *Photogrammetric Engineering & Remote Sensing* 71(5): 595–602. http://www.ingentaconnect.com/content/asprs/pers/2005/00000071/00000005/art00002?crawler=true.

Weiss, Marie, and Frederic Baret. 2017. "Using 3D Point Clouds Derived from UAV RGB Imagery to Describe Vineyard 3D Macro- Structure." *Remote Sensing* 9(2).

Wijesingha, Jayan, Thomas Moeckel, Frank Hensgen, and Michael Wachendorf. 2019. "Evaluation of 3D Point Cloud-Based Models for the Prediction of Grassland Biomass." *Int J Appl Earth Obs Geoinformation* 78(June): 352–59. https://doi.org/10.1016/j.jag.2018.10.006.

Yu, Kang et al. 2013. "Remotely Detecting Canopy Nitrogen Concentration and Uptake of Paddy Rice in the Northeast China Plain." *ISPRS Journal of Photogrammetry and Remote Sensing* 78: 102–15. http://dx.doi.org/10.1016/j.isprsjprs.2013.01.008 (April 25, 2019).

Yu, Neil et al. 2016. "Development of Methods to Improve Soybean Yield Estimation and Predict Plant Maturity with an Unmanned Aerial Vehicle Based Platform." *Remote Sensing of Environment* 187: 91–101.

Yue, Jibo et al. 2017. “Estimation of Winter Wheat Above-Ground Biomass Using Unmanned Aerial Vehicle-Based Snapshot Hyperspectral Sensor and Crop Height Improved Models.” *Remote Sensing* 9(7).

Yue, Jibo, Haikuan Feng, Xiuliang Jin, et al. 2018. “A Comparison of Crop Parameters Estimation Using Images from UAV-Mounted Snapshot Hyperspectral Sensor and High-Definition Digital Camera.” *Remote Sensing* 10(7): 1138. http://www.mdpi.com/2072-4292/10/7/1138/pdf.

Yue, Jibo, Haikuan Feng, Guijun Yang, and Zhenhai Li. 2018. “A Comparison of Regression Techniques for Estimation of Above-Ground Winter Wheat Biomass Using Near-Surface Spectroscopy.” *Remote Sensing* 10(2): 66. http://www.mdpi.com/2072-4292/10/1/66.

Zhang, Chunhua, and John M Kovacs. 2012. “The Application of Small Unmanned Aerial Systems for Precision Agriculture: A Review.” *Precision Agriculture* 13(6): 693–712.

Zhang, F., and G. Zhou. 2019. “Estimation of Vegetation Water Content Using Hyperspectral Vegetation Indices: A Comparison of Crop Water Indicators in Response to Water Stress Treatments for Summer Maize.” *BMC Ecology* 19(1): 18. https://bmcecol.biomedcentral.com/articles/10.1186/s12898-019-0233-0 (May 23, 2019).

Chapter 9

Gallery

Lablab, finger millet and maize grown in rainfed and irrigated experimental fields of monsoon season (June to October)

Lablab

Finger millet

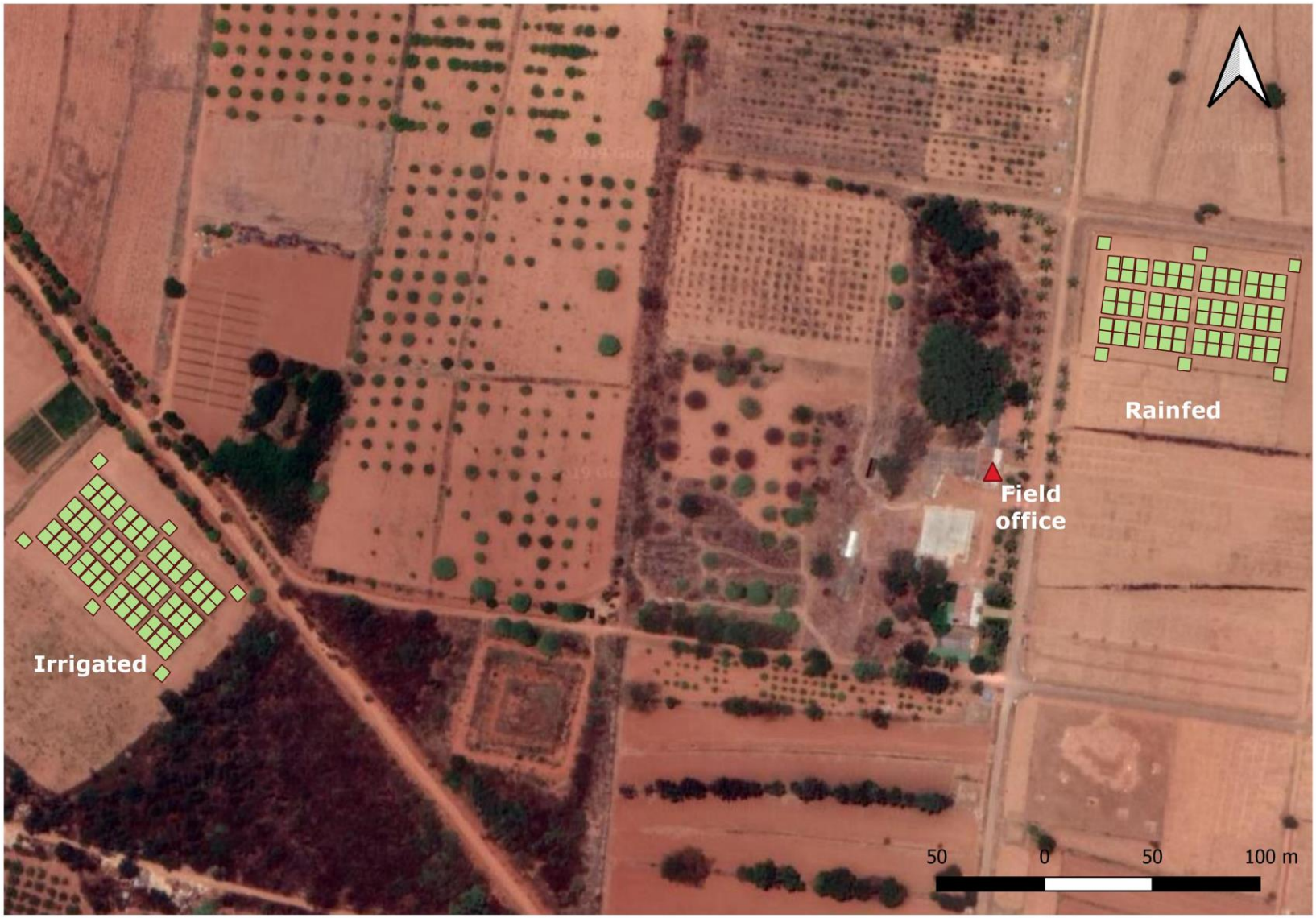
Maize

Eggplant, tomato and cabbage grown in irrigated experimental field of dry season (January to May)

Eggplant

Tomato

Cabbage



Google satellite layer image showing irrigated, rainfed experiments and field office at GKVK campus of UAS, Bengaluru.